

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 12, 2006, 19:48:01 ; Search time 195 Seconds
(without alignments)
1034.230 Million cell updates/sec

Title: US-09-967-237B-2

Perfect score: 2424
Sequence: 1 MAPLCSPWLLPAPAG.....RRGTGGSYRPAEVAETGA 459

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A Geneseq_21.*
1: Geneseqp1980s.*
2: Geneseqp1990s.*
3: Geneseqp2000s.*
4: Geneseqp2001s.*
5: Geneseqp2002s.*
6: Geneseqp2003as.*
7: Geneseqp2003bs.*
8: Geneseqp2004s.*
9: Geneseqp2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2424	100.0	459	2 AAR8058	Aar8058 Protein e
2	2424	100.0	459	3 AAY53228	Aay53228 Human MN
3	2424	100.0	459	3 AAB03005	Aab03005 Human MN
4	2424	100.0	459	5 AAE17175	Aae17175 Human RCC
5	2424	100.0	459	6 ABR58596	Abr58596 Human can
6	2424	100.0	459	6 ABP97744	Abp97744 Amino aci
7	2424	100.0	459	6 ABU56656	Abu56656 Lung canc
8	2424	100.0	459	6 ABU56408	Abu56408 Lung canc
9	2424	100.0	459	7 ADN38988	Adn38988 Cancer/an
10	2424	100.0	459	7 ADN39921	Adn39921 Cancer/an
11	2424	100.0	459	8 ADG31414	Adg31414 Human MN
12	2424	100.0	459	8 ADK41804	Adk41804 Human MN
13	2424	100.0	459	8 ADL70156	Adl70156 Human Car
14	2424	100.0	459	8 ADK15606	Adk15606 Human G25
15	2424	100.0	459	8 ADQ17644	Adq17644 Human sof
16	2424	100.0	459	8 ADP53999	Adp53999 Human car
17	2424	100.0	459	8 ADQ29712	Adq29712 Human col
18	2424	100.0	459	8 ADS88409	Ads88409 Human pro
19	2424	100.0	459	8 ADT36534	Adt36534 Renal cel
20	2424	100.0	459	9 ADZ64595	Adz64595 Human MN/
21	2419	99.8	610	4 AAB82848	Aab82848 Kidney ca
22	2416	99.7	459	9 AEA00140	Aea00140 Human TAT
23	2416	99.7	459	9 AEA00660	Aea00660 Human TAT
24	2007	82.8	377	3 AAY53245	Aay53245 MN protei

25	2007	82.8	377	3 AAB03021	Aab03021 Mature hu
26	2007	82.8	377	8 ADG31451	Adg31451 Extracell
27	2007	82.8	377	9 ADZ64600	Adz64600 Human MN/
28	1844.5	76.1	362	8 ADL70160	Adl70160 Human car
29	1600	66.0	325	8 ADL70182	Adl70182 Glu-tagge
30	1596	65.8	437	8 ADG31476	Adg31476 Murine ca
31	1576	65.0	325	8 ADL70164	Adl70164 Glu-tagge
32	1478	61.0	277	8 ADG31456	Adg31456 Carbonic
33	1398.5	57.7	429	2 AAR41746	Aar41746 MN protei
34	1370	56.5	257	3 AAY53241	Aay53241 MN protei
35	1370	56.5	257	3 AAB03018	Aab03018 Human MN
36	1368	56.4	257	9 ADZ64598	Adz64598 Human agg
37	1364	56.3	256	2 AAR97235	Aar97235 MuTu puta
38	1340	55.3	358	8 ADG31478	Adg31478 N-termina
39	1210	49.9	232	8 ADG31469	Adg31469 Antigenic
40	1108	45.7	258	8 ADG31480	Adg31480 Murine MN
41	763	31.5	179	8 ADL70158	Adl70158 Human car
42	562	23.2	337	3 AAY99460	Aay99460 Human PRO
43	562	23.2	337	4 AAB66209	Aab66209 Protein O
44	562	23.2	337	4 AAU29157	Aau29157 Human PRO
45	562	23.2	337	4 AAB87562	Aab87562 Human PRO

ALIGNMENTS

RESULT 1

AAR88058
ID AAR88058 standard; protein; 459 AA.

XX AAR88058;
AC
XX
DT 25-MAR-2003 (revised)
DT 25-JUL-1996 (first entry)
XX
DE Protein encoded by MuTu putative oncogene MN.

XX MuTu; endogenous; cellular component; MN; HeLa cell; diagnosis;
XX Lymphocytic choriomeningitis virus; LCMV; putative oncogene; treatment;
XX neoplastic; pre-neoplastic; disease; antitense therapy; antibody;
XX vaccine; vertebrate; immunisation; carbonic anhydrase.

OS Homo sapiens.

XX	Key	Location/Qualifiers
FT	Peptide	1..37
FT	Peptide	/label= sig_peptide
FT	Peptide	/note= "anti-MN antibody epitope"
FT	Region	38..135
FT	Peptide	/note= "region homologous to collagen alpha 1 chain"
FT	Peptide	55..60
FT	Peptide	/note= "anti-MN antibody epitope"
FT	Peptide	62..67
FT	Peptide	/note= "anti-MN antibody epitope"
FT	Peptide	68..91
FT	Peptide	/note= "anti-MN antibody epitope"
FT	Peptide	127..147
FT	Peptide	/note= "anti-MN antibody epitope"
FT	Domain	136..391
FT	Peptide	/note= "carbonic anhydrase domain"
FT	Peptide	279..291
FT	Peptide	/note= "anti-MN antibody epitope"
FT	Region	414..433
FT	Region	/note= "intracellular transmembrane region"
FT	Region	434..459
FT	Region	/note= "intracellular C-terminus"
FT	Peptide	435..450
FT	Peptide	/note= "anti-MN antibody epitope"

XX W05934650-A2.

XX 21-DEC-1995.

```
XX 15-JUN-1995; 95WO-US0007628.
XX 15-JUN-1994; 94US-00260190.
PR 07-JUN-1995; 95US-00477504.
PR 07-JUN-1995; 95US-00481658.
PR 07-JUN-1995; 95US-00485049.
PR 07-JUN-1995; 95US-00485049.
PR 07-JUN-1995; 95US-00485862.
PR 07-JUN-1995; 95US-00485863.
PR 07-JUN-1995; 95US-00486756.
PR 07-JUN-1995; 95US-00487077.
XX (CIBA ) CIBA CORNING DIAGNOSTICS CORP.
PA (VIRO-) INST VIROLOGY.
XX
XX Zavada J, Pastorekova S, Pastorek J;
DR WPI; 1996-049679/05.
DR N-PSDB; AAT09186.
XX
XX MN gene, protein and nucleic acid fragments - used as primers and probes
PT in the detection of MN antigens and antibodies, and in the treatment of
PT (pre)neoplastic disease.
XX
XX Claim 12; Fig 1; 102pp; English.
XX
XX The present sequence is encoded by the full length MuTu endogenous
CC cellular component, MN, cDNA clone, which was isolated from lymphocytic
CC choriomeningitis virus (LCMV) infected Hela cells. Persistent LCMV, the
CC exogenous MuTu transmissible agent (MX), infection increases the
CC expression level of the MN gene. MN is a putative oncogene, and can
CC therefore be used in the development of prods. for the diagnosis and
CC treatment of neoplastic (NP), or pre-NP diseases. NP diseases can be
CC treated using DNA antisense to MN transcribed mRNA, anti-MN protein
CC antibodies can be used for the diagnosis NP or pre-NP diseases and a
CC vaccine contg. immunogenic amounts of the MN protein can be used to
CC immunise a vertebrate against a NP disease associated with MN antigen
CC expression. (Updated on 25-MAR-2003 to correct PR field.)
XX
XX Sequence 459 AA;
XX
XX Query Match 100.0%; Score 2424; DB 2; Length 459;
XX Best Local Similarity 100.0%; Pred. No. 9,9e-183;
XX Matches 459; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 MAPLCPSPWMLPLIIPAPAGLTVQLLSLLMLPVPQRLPRMQEDSPILGGSSGGDDPL 60
DB 1 MAPLCPSPWMLPLIIPAPAGLTVQLLSLLMLPVPQRLPRMQEDSPILGGSSGGDDPL 60
QY 61 GEEDLPSEDSPREDDPPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLP 120
DB 61 GEEDLPSEDSPREDDPPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLP 120
QY 121 DPOEPONNAHRKEDGDDQSHWRVGGDPMPRVSPACAGRFQSPVDTRPOLAFCPALRPL 180
DB 121 DPOEPONNAHRKEDGDDQSHWRVGGDPMPRVSPACAGRFQSPVDTRPOLAFCPALRPL 180
QY 181 ELLGFGQLPPLPELRLNNGSHVQLTLPGLMALGFGREYRALQLHLHWAAGRPGSEHT 240
DB 181 ELLGFGQLPPLPELRLNNGSHVQLTLPGLMALGFGREYRALQLHLHWAAGRPGSEHT 240
QY 241 VEGHRRFAEIHVVHLSTAFARVDEALGRPGGLAVLAFLAEEGPEENSAYEQLLSRLEETA 300
DB 241 VEGHRRFAEIHVVHLSTAFARVDEALGRPGGLAVLAFLAEEGPEENSAYEQLLSRLEETA 300
QY 301 EGSEGTQVCLDLSALLPSPDFSYFQYEGSLTTPPCAQGVITVFNQTVMLSAKQLHTLS 360
DB 301 EGSEGTQVCLDLSALLPSPDFSYFQYEGSLTTPPCAQGVITVFNQTVMLSAKQLHTLS 360
QY 361 DTLLWGFQSDRLQNLNFRATQPLNGRVIEASFPAGVDSSPRAAEFVQLNSCLAGDILALVF 420
DB 361 DTLLWGFQSDRLQNLNFRATQPLNGRVIEASFPAGVDSSPRAAEFVQLNSCLAGDILALVF 420
```

```
QY 421 GLLFAVTSVAFLVQMRQHRRTKGGVSYRPAEVAETGA 459
DB 421 GLLFAVTSVAFLVQMRQHRRTKGGVSYRPAEVAETGA 459
RESULT 2
AAY53228
ID AAY53228 standard; protein; 459 AA.
XX
XX AAY53228;
XX
XX 16-JUN-2000 (first entry)
XX
XX Human MN protein SEQ ID NO:2.
XX
XX Human; MN protein; MN gene; oncogene; carbonic anhydrase; tumour;
KW oncogenesis; diagnosis; neoplastic disease; cancer; carcinoma;
KW MN/CA IX isoenzyme.
XX
XX Homo sapiens.
XX
XX US6027887-A.
XX
XX 22-FEB-2000.
XX
XX 24-JAN-1997; 97US-00787739.
XX
XX 21-OCT-1992; 92US-00964589.
PR 30-DEC-1993; 93US-00177093.
PR 15-JUN-1994; 94US-00260190.
PR 07-JUN-1995; 95US-00477504.
PR 07-JUN-1995; 95US-00481658.
PR 07-JUN-1995; 95US-00485049.
PR 07-JUN-1995; 95US-00485862.
PR 07-JUN-1995; 95US-00485863.
PR 07-JUN-1995; 95US-00486756.
PR 07-JUN-1995; 95US-00487077.
XX
XX (SLSC-) SLOVAK ACAD SCI INST VIROLOGY.
XX
XX Pastorek J, Zavada J, Pastorekova S;
XX WPI; 2000-194827/17.
XX N-PSDB; AAA16540.
XX
XX Nucleic acid based assay for diagnosing a wide variety of
PT preneoplastic/neoplastic disease comprises screening for the presence of
PT abnormal MN gene expression in a vertebrate.
XX
XX Disclosure; Fig 1; 87pp; English.
XX
XX The present invention describes a method of screening for
CC preneoplastic/neoplastic disease. The method comprises: (1) determining
CC whether abnormal MN gene expression is present in a vertebrate; and (2)
CC if abnormal MN gene expression is determined to be present in the
CC vertebrate, determining that the vertebrate has a significant risk of
CC having preneoplastic/neoplastic disease. The MN gene is an oncogene and
CC encodes an MN protein (also referred to as MN/CA IX isoenzyme). The MN
CC protein is a tumour associated carbonic anhydrase isoenzyme. The method
CC is used for detecting a wide variety of preneoplastic/neoplastic diseases
CC in a vertebrate, preferably a human. The disease detected is mammary,
CC bladder, renal, urinary tract, ovarian, uterine, cervical, endometrial,
CC vaginal, vulval, prostate, liver, lung, skin, thyroid, pancreatic,
CC testicular, brain, head and neck, mesodermal, gallbladder, rectal,
CC duodenal, jejunal, ileal, gastric, pancreatic duct, liver duct, gastric
CC mucosa, gallbladder epithelium, small intestinal mucosa, colorectal
CC mucosa, pancreatic duct epithelium or liver duct epithelium
CC preneoplastic/neoplastic disease. AAA16540 to AAA16617 and AAY53228 to
CC AAY53245 represent sequences used in the exemplification of the present
CC invention
XX
XX Sequence 459 AA;
```

RESULT 4

KW chronic obstructive pulmonary disease; hypersensitivity pneumonitis;
KW interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.
XX Unidentified.
OS WO200286443-A2.
XX 31-OCT-2002.
XX 18-APR-2002; 2002WO-US012476.
XX 18-APR-2001; 2001US-0284770P.
PR 10-MAY-2001; 2001US-0290492P.
PR 09-NOV-2001; 2001US-0339245P.
PR 13-NOV-2001; 2001US-0350666P.
PR 29-NOV-2001; 2001US-0334370P.
PR 12-APR-2002; 2002US-0372246P.
XX (EOSB-) EOS BIOTECHNOLOGY INC.
XX Aziz N, Murray R;
PI WPI; 2003-093161/08.
XX N-PSDB; ABX76385.
XX Detecting a lung cancer-associated transcript in a cell from a patient
PT for treating lung cancer, by contacting a biological sample from the
PT patient with a polynucleotide that exhibits increased or decreased
PT expression in lung cancer.
PS Claim 27; Page 380; 453pp; English.
XX The invention relates to a method for detecting a lung cancer-associated
CC transcript in a cell from a patient, comprising contacting a biological
CC sample from the patient with a polynucleotide that selectively hybridises
CC to a sequence that is at least 80 % identical to a gene that exhibits
CC increased or decreased expression in lung cancer samples. Lung cancer-
CC associated polynucleotides and polypeptides are used for identifying a
CC compound that modulates a lung cancer-associated polypeptide, for
CC inhibiting proliferation of a lung cancer-associated cell to treat lung
CC cancer in a patient and for treating a mammal having lung cancer by
CC administering a modulatory compound identified. The methods are useful
CC for treating lung cancer, such as small cell lung cancer, non-small cell
CC lung cancer or other benign or precancerous lesions, e.g. atelectasis,
CC emphysema, bronchitis, chronic obstructive pulmonary disease, fibrosis,
CC hypersensitivity pneumonitis, interstitial pulmonary fibrosis, asthma and
CC bronchiectasis. The genes, polynucleotides and polypeptides are useful
CC for diagnostic purposes and as targets for screening for therapeutic
CC compounds that modulate lung cancer, such as antibodies. Sequences
CC ABU56408-ABU56745 represent lung cancer-associated polypeptides of the
CC invention
XX Sequence 459 AA;
SQ
Query Match 100.0%; Score 2424; DB 6; Length 459;
Best Local Similarity 100.0%; Pred. No. 9.9e-183;
Matches 459; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MAPLCSPWMLPLIIPAPAGLTWQLLSLLMLMPVHPQRLPRMOEDSLPGGSGSGDDPL 60
DB 1 MAPLCSPWMLPLIIPAPAGLTWQLLSLLMLMPVHPQRLPRMOEDSLPGGSGSGDDPL 60
QY 61 GEEDLPSEEDSPREEDPPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLP 120
DB 61 GEEDLPSEEDSPREEDPPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLP 120
QY 121 DPOEPQNNARHDEKGDQSHWRVGGDPWPVRVSPACAGRFQSPVDTRPOLAFCPALRPL 180
DB 121 DPOEPQNNARHDEKGDQSHWRVGGDPWPVRVSPACAGRFQSPVDTRPOLAFCPALRPL 180
QY 181 ELLGFGOLPPLPELRLRNNGHSVOLTLPGLLEWALGFGREYRALQLHLHWGAAGRPGSEHT 240
DB 181 ELLGFGOLPPLPELRLRNNGHSVOLTLPGLLEWALGFGREYRALQLHLHWGAAGRPGSEHT 240

QY 241 VEGHRRPAAIHVHILSTAFARVDEALGRPGGLAVLAALFLEEGPEENSAYEQLLSRLEEIA 300
DB 241 VEGHRRPAAIHVHILSTAFARVDEALGRPGGLAVLAALFLEEGPEENSAYEQLLSRLEEIA 300
QY 301 EEGSETQVPGLDISALLPSDFSRYFOYEGSLTTPCAQGVITWVFNQTVMLSAKQLHTLS 360
DB 301 EEGSETQVPGLDISALLPSDFSRYFOYEGSLTTPCAQGVITWVFNQTVMLSAKQLHTLS 360
QY 361 DTLWGPDSRLQLNFRATOPNLNGRVIEASFAGVDSPPRAASPVLQNSCLAAGDILALVF 420
DB 361 DTLWGPDSRLQLNFRATOPNLNGRVIEASFAGVDSPPRAASPVLQNSCLAAGDILALVF 420
QY 421 GLLFAVTSVAFVLQVRRQRHRTKGVSYRPAEVAETGA 459
DB 421 GLLFAVTSVAFVLQVRRQRHRTKGVSYRPAEVAETGA 459
RESULT 8
ABU56408
ID ABU56408 standard; protein; 459 AA.
XX
AC ABU56408;
XX
DT 02-APR-2003 (first entry)
XX Lung cancer-associated polypeptide #1.
DE
XX Lung cancer-associated polypeptide; cytostatic; emphysema;
KW antiinflammatory; antiasthmatic; non-small cell lung cancer; atelectasis;
KW small cell lung cancer; benign lesion; precancerous lesion; bronchitis;
KW chronic obstructive pulmonary disease; hypersensitivity pneumonitis;
KW interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.
XX Unidentified.
OS WO200286443-A2.
XX 31-OCT-2002.
XX 18-APR-2002; 2002WO-US012476.
PR 18-APR-2001; 2001US-0284770P.
PR 10-MAY-2001; 2001US-0290492P.
PR 09-NOV-2001; 2001US-0339245P.
PR 13-NOV-2001; 2001US-0350666P.
PR 29-NOV-2001; 2001US-0334370P.
PR 12-APR-2002; 2002US-0372246P.
XX (EOSB-) EOS BIOTECHNOLOGY INC.
PA Aziz N, Murray R;
XX
PI WPI; 2003-093161/08.
XX N-PSDB; ABX76124.
XX Detecting a lung cancer-associated transcript in a cell from a patient
PT for treating lung cancer, by contacting a biological sample from the
PT patient with a polynucleotide that exhibits increased or decreased
PT expression in lung cancer.
XX Claim 27; Page 189; 453pp; English.
XX The invention relates to a method for detecting a lung cancer-associated
CC transcript in a cell from a patient, comprising contacting a biological
CC sample from the patient with a polynucleotide that selectively hybridises
CC to a sequence that is at least 80 % identical to a gene that exhibits
CC increased or decreased expression in lung cancer samples. Lung cancer-
CC associated polynucleotides and polypeptides are used for identifying a
CC compound that modulates a lung cancer-associated polypeptide, for
CC inhibiting proliferation of a lung cancer-associated cell to treat lung
CC cancer in a patient and for treating a mammal having lung cancer by
CC administering a modulatory compound identified. The methods are useful

CC for treating lung cancer, such as small cell lung cancer, non-small cell
CC lung cancer or other benign or precancerous lesions, e.g. atelectasis,
CC emphysema, bronchitis, chronic obstructive pulmonary disease, fibrosis, and
CC hypersensitivity pneumonitis, interstitial pulmonary fibrosis, asthma and
CC bronchiectasis. The genes, polynucleotides and polypeptides are useful
CC for diagnostic purposes and as targets for screening for therapeutic
CC compounds that modulate lung cancer, such as antibodies. Sequences
CC ABU56408-ABU56745 represent lung cancer-associated polypeptides of the
CC invention
XX
SQ Sequence 459 AA;

Query Match 100.0%; Score 2424; DB 6; Length 459;
Best Local Similarity 100.0%; Pred. No. 9.9e-183;
Matches 459; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MAPLCPSPWLPPLLPAPAPGLTVQLLSLLLMVHPQRLPRMQEDSPGSGSGEDDPL 60
Db 1 MAPLCPSPWLPPLLPAPAPGLTVQLLSLLLMVHPQRLPRMQEDSPGSGSGEDDPL 60
Qy 61 GEEDLPSEDSPREDDPGEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLP 120
Db 61 GEEDLPSEDSPREDDPGEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLP 120
Qy 121 DPQEPQNNNAHRDKGDDQSHWRYGDDPPWPRVSPACAGRFQSPVDIRPQLAFCPALRPL 180
Db 121 DPQEPQNNNAHRDKGDDQSHWRYGDDPPWPRVSPACAGRFQSPVDIRPQLAFCPALRPL 180
Qy 181 ELLGFQLPPLPELRLRNNHSHVQLTLPGLMALGPGREYRALQHLHFWGAAGRPGSEHT 240
Db 181 ELLGFQLPPLPELRLRNNHSHVQLTLPGLMALGPGREYRALQHLHFWGAAGRPGSEHT 240
Qy 241 VEGHRFPAAIHVHLSTAFARVDEALGRPGGLAVLAFLRPGPENSAYEQLLSRLEIA 300
Db 241 VEGHRFPAAIHVHLSTAFARVDEALGRPGGLAVLAFLRPGPENSAYEQLLSRLEIA 300
Qy 301 EGSETQVPGIDISALLPDSFRYFQVSGSLTTPCAQGVITWVFNQTVMLSAKQLHTLS 360
Db 301 EGSETQVPGIDISALLPDSFRYFQVSGSLTTPCAQGVITWVFNQTVMLSAKQLHTLS 360
Qy 361 DTLWPGDSRLQLNFRATQPLNGRVIEASFAGVDSSPRAAEPVQLNSCLAAGDILALVF 420
Db 361 DTLWPGDSRLQLNFRATQPLNGRVIEASFAGVDSSPRAAEPVQLNSCLAAGDILALVF 420
Qy 421 GLLFAVTSVAFVQMRQRHRTGKGVSYPRAEVAETGA 459
Db 421 GLLFAVTSVAFVQMRQRHRTGKGVSYPRAEVAETGA 459

RESULT 9
ADN38988
ID ADN38988 standard; protein; 459 AA.
XX
AC ADN38988;
XX
DT 17-JUN-2004 (first entry)
XX
DE Cancer/angiogenesis/fibrosis-related polypeptide, SEQ ID NO:306.
XX
KW Human; differential expression; cancer; angiogenic disorder;
KW fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;
KW inflammatory disease; autoimmune disease;
KW retinal neovascularization syndrome; scarring; uterine fibroid;
KW detection; diagnosis; prognosis; drug screening; drug targeting;
KW wound healing; contraception; cytostatic; cardiant; immunomodulatory;
KW vulnery; gene therapy; vaccine.
XX
OS Homo sapiens.
XX
PN WO2003042661-A2.
XX
PD 22-MAY-2003.
XX

PF 13-NOV-2002; 2002WO-US036810.
XX
PR 13-NOV-2001; 2001US-0350666P.
PR 21-NOV-2001; 2001US-0332464P.
PR 29-NOV-2001; 2001US-0334393P.
PR 03-DEC-2001; 2001US-0335394P.
PR 14-DEC-2001; 2001US-0340376P.
PR 08-JAN-2002; 2002US-0347211P.
PR 10-JAN-2002; 2002US-0347349P.
PR 08-FEB-2002; 2002US-035250P.
PR 13-FEB-2002; 2002US-0356714P.
PR 20-FEB-2002; 2002US-0359077P.
PR 29-MAR-2002; 2002US-0368809P.
PR 04-APR-2002; 2002US-0370110P.
PR 12-APR-2002; 2002US-0372246P.
PR 05-JUN-2002; 2002US-0386614P.
PR 16-JUL-2002; 2002US-0396839P.
PR 22-JUL-2002; 2002US-0397757P.
PR 22-JUL-2002; 2002US-0397845P.
PR 09-SEP-2002; 2002US-0409450P.
XX
PA (EOSB-) EOS BIOTECHNOLOGY INC.
XX
XX Afar D, Aziz N, Ginsburg WM, Gish KC, Glynn R, Hevezi PA;
PI Mack DH, Murray R, Watson SR, Wilson KE, Zlotnik A;
XX
XX WPI; 2003-468649/44.
DR N-PSDB; ADN38987.
XX
XX Determining the presence or absence of a pathological cell in a patient,
XX useful for diagnosing, prognosing or treating cancer, comprises detecting
XX a nucleic acid in a biological sample.
XX
XX Claim 12; SEQ ID NO 306; 1385pp; English.
XX
XX The invention relates to nucleic acids and proteins (ADN38683-ADN40064)
XX whose expression is upregulated or downregulated in specific cancers or
XX other diseases such as angiogenic or fibrotic disorders, and to methods
XX of determining the presence or absence of a pathological cell in a
XX patient by detecting a nucleic acid at least 80% identical to those of
XX the invention or by detecting a polypeptide of the invention. The
XX invention also relates to expression vectors and host cells comprising a
XX nucleic acid of the invention; antibodies which specifically bind a
XX polypeptide of the invention; use of such antibodies for drug targeting;
XX and methods of screening for modulators of activity or expression of the
XX polypeptides and nucleic acids. The nucleic acids, polypeptides,
XX antibodies and methods are useful for diagnosing, prognosing and treating
XX cancer and other conditions such as psoriasis, ischaemia, heart disease,
XX atherosclerosis, inflammatory diseases, autoimmune diseases, retinal
XX neovascularization syndromes, scarring and uterine fibroids. They may
XX also be useful in wound healing and in contraception. The present
XX sequence represents a polypeptide of the invention.
XX
SQ Sequence 459 AA;
Query Match 100.0%; Score 2424; DB 7; Length 459;
Best Local Similarity 100.0%; Pred. No. 9.9e-183;
Matches 459; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MAPLCPSPWLPPLLPAPAPGLTVQLLSLLLMVHPQRLPRMQEDSPGSGSGEDDPL 60
Db 1 MAPLCPSPWLPPLLPAPAPGLTVQLLSLLLMVHPQRLPRMQEDSPGSGSGEDDPL 60
Qy 61 GEEDLPSEDSPREDDPGEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLP 120
Db 61 GEEDLPSEDSPREDDPGEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLP 120
Qy 121 DPQEPQNNNAHRDKGDDQSHWRYGDDPPWPRVSPACAGRFQSPVDIRPQLAFCPALRPL 180
Db 121 DPQEPQNNNAHRDKGDDQSHWRYGDDPPWPRVSPACAGRFQSPVDIRPQLAFCPALRPL 180
Qy 181 ELLGFQLPPLPELRLRNNHSHVQLTLPGLMALGPGREYRALQHLHFWGAAGRPGSEHT 240
XX

KW CA9; MN/CA9; carbonic anhydrase IX; CA IX; MN/G250; neoplastic tumour;
 KW cytosolic; cancer; tumour therapy; anti-tumour vaccine.
 XX Homo sapiens.
 XX WO2003100029-A2.
 XX PD 04-DEC-2003.
 XX XX 22-FEB-2003; 2003WO-US005136.
 XX PR 23-MAY-2002; 2002US-0383068P.
 XX PR 05-DEC-2002; 2002US-0431499P.
 XX XX (FARB) BAYER CORP.
 XX PA (VIRO-) INST VIROLOGY.
 XX PI Pastorek J, Pastorekova S, Zatovicova M, Zavada J, Ortova Gut M;
 XX DR WPI; 2004-035136/03.
 XX DR N-PSDB; ADG31413, ADG31415.
 XX XX New monoclonal antibody generated from MN/CA IX-deficient mice, where the
 XX antibody binds specifically to human tumor-associated cell adhesion
 PT protein MN/CA IX or polypeptide, useful for diagnosing, prognosing or
 PT treating cancer.
 XX XX Disclosure; SEQ ID NO 2; 156pp; English.
 XX CC This invention relates to a novel monoclonal antibody identified as the
 CC MN/CA IX specific antibody prepared in knockout mice (CA IX deficient
 CC mice). Specifically, this antibody is directed towards the MN gene, a
 CC cellular oncogene known alternatively as carbonic anhydrase 9, CA9 or
 CC MN/CA9, which encodes the MN protein that is also known as the MN/CA IX
 CC isoenzyme, carbonic anhydrase IX, CA IX or the MN/G250 protein. The
 CC present invention describes the generation of this monoclonal antibody,
 CC and immunoreactive fragments thereof, which are directed against non-
 CC immunodominant epitopes on the CA IX extracellular domain. As such, this
 CC antibody can be useful diagnostically as a marker for preneoplastic/
 CC neoplastic tumours, immunodetection methods and immunotargeting
 CC approaches. Accordingly, compositions exhibit cytostatic activity and are
 CC useful in the diagnosis, prognosis and treatment of various cancers
 CC including breast, bladder or lung cancer, in tumour therapy and in anti-
 CC tumour vaccination. This polypeptide sequence is the human MN protein of
 CC the invention.
 XX XX Sequence 459 AA;
 XX Query Match 100.0%; Score 2424; DB 8; Length 459;
 XX Best Local Similarity 100.0%; Pred. No. 9.9e-183;
 XX Matches 459; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MAPLCPSWLPPLIPAPAGPLTVQLLSILLMPVHPQRLPRMORSDSLGGSGGDDPL 60
 DB 1 MAPLCPSWLPPLIPAPAGPLTVQLLSILLMPVHPQRLPRMORSDSLGGSGGDDPL 60
 QY 61 GEEDLPSEDSPREDDPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLP 120
 DB 61 GEEDLPSEDSPREDDPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLP 120
 QY 121 DPQEPQNNARHDKGDDQSHWRYGDDPWPRVSPACAGRFQSPVDIRPQLAAFCPALRPL 180
 DB 121 DPQEPQNNARHDKGDDQSHWRYGDDPWPRVSPACAGRFQSPVDIRPQLAAFCPALRPL 180
 QY 181 ELLGQLPPLPELRLNNGHVSQVLTLPGLMALGPGEYRALQLHLHWAAGRPGSEHT 240
 DB 181 ELLGQLPPLPELRLNNGHVSQVLTLPGLMALGPGEYRALQLHLHWAAGRPGSEHT 240
 QY 241 VEGHFRPAEIHVVHLSFAFARVDEALGRPGGLAVLAAFLFEGPEENSAYEQLLSLREIA 300
 DB 241 VEGHFRPAEIHVVHLSFAFARVDEALGRPGGLAVLAAFLFEGPEENSAYEQLLSLREIA 300
 QY 301 EEGSETQVPGLDISALLPSDFSRFYQYEGSLTTPPCAQGVITVFNQTVMLSAKQLHTLS.360.

Db 301 EEGSETQVPGLDISALLPSDFSRFYQYEGSLTTPPCAQGVITVFNQTVMLSAKQLHTLS 360
 QY 361 DTLWGPGDSRLQNLFRATOPNLNGRVIEASFPAGVDSSPRAAEPVOLNSCLAAGDILALVF 420
 Db 361 DTLWGPGDSRLQNLFRATOPNLNGRVIEASFPAGVDSSPRAAEPVOLNSCLAAGDILALVF 420
 QY 421 GLLFVNTSVAFLVQMRQRHRRGTGGVSYRPAEVAETGA 459
 Db 421 GLLFVNTSVAFLVQMRQRHRRGTGGVSYRPAEVAETGA 459
 RESULT 12
 ADK41804
 ID ADK41804 standard; protein; 459 AA.
 XX AC ADK41804;
 XX DT 06-MAY-2004 (first entry)
 XX DE Human MN protein sequence SeqID2.
 XX KW carbonic anhydrase IX; CA IX; precancerous cell; MN; cancerous cell;
 KW human; vertebrae; cytostatic; vaccine; gene therapy;
 KW renal cell carcinoma; breast cancer; colorectal cancer.
 XX OS Homo sapiens.
 XX FN WO2004005348-A1.
 XX PD 15-JAN-2004.
 XX XX 22-FEB-2003; 2003WO-US005137.
 XX PR 23-MAY-2002; 2002US-0383068P.
 XX PR 05-DEC-2002; 2002US-0431499P.
 XX XX (FARB) BAYER CORP.
 XX XX (VIRO-) INST VIROLOGY.
 XX PI Zavada J, Pastorekova S, Pastorek J, Zavadova Z;
 XX WPI; 2004-083500/08.
 XX DR N-PSDB; ADK41803, ADK41805.
 XX XX New soluble form of the carbonic anhydrase IX (CA IX) protein for
 XX screening, diagnosing or prognosing diseases associated with abnormal
 XX expression of CA IX protein, e.g. renal cell carcinoma, breast cancer or
 XX colorectal cancer.
 XX PT Disclosure; SEQ ID NO 2; 159pp; English.
 XX CC This invention relates to a novel soluble form of the carbonic anhydrase
 CC IX (CA IX) (or MN) protein or CA IX polypeptide which is released from
 CC precancerous and/or cancerous cells of a vertebrate into a body fluid.
 CC The invention may be useful for the development of compounds with a
 CC cytostatic activity or a vaccine whilst the disclosed sequences may be
 CC used for gene therapy. The protein and method are useful for screening,
 CC diagnosing or prognosing diseases associated with abnormal expression of
 CC carbonic anhydrase IX protein, such as precancerous and cancerous
 CC diseases like renal cell carcinoma, breast cancer or colorectal cancer.
 CC The monoclonal antibody may also be used for treating or preventing
 CC precancerous and cancerous diseases. The present sequence is that of the
 CC human MN protein of the invention.
 XX SQ Sequence 459 AA;
 XX Query Match 100.0%; Score 2424; DB 8; Length 459;
 XX Best Local Similarity 100.0%; Pred. No. 9.9e-183;
 XX Matches 459; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MAPLCPSWLPPLIPAPAGPLTVQLLSILLMPVHPQRLPRMORSDSLGGSGGDDPL 60
 QY 1 MAPLCPSWLPPLIPAPAGPLTVQLLSILLMPVHPQRLPRMORSDSLGGSGGDDPL 60

—

Result No.	Score	Query		DB	ID	Description
		Match	%			
1	2424	100.0	459	2	I38013	p54/58N - human
2	526	21.7	307	2	A29993	carbonate dehydrat
3	511	21.1	319	2	S71877	carbonate dehydrat
4	492	20.3	308	1	CRU6	carbonate dehydrat
5	425	17.5	303	2	S12867	carbonate dehydrat
6	412	17.0	290	2	JN0576	carbonate dehydrat
7	409	16.9	264	1	CRU7	carbonate dehydrat
8	402.5	16.6	260	2	T08463	carbonate dehydrat
9	391	16.1	259	1	CRB02	carbonate dehydrat
10	380.5	15.7	260	1	CRMS2	carbonate dehydrat
11	379.5	15.7	260	1	CRHU2	carbonate dehydrat
12	379	15.6	259	1	CRSH2	carbonate dehydrat
13	374.5	15.4	259	1	CRRB2	carbonate dehydrat
14	368.5	15.2	260	2	JH0527	carbonate dehydrat
15	366.5	15.1	260	1	CRM01R	carbonate dehydrat
16	364.5	15.0	260	2	JC2580	carbonate dehydrat
17	363	15.0	261	2	A26344	carbonate dehydrat
18	362.5	15.0	261	2	JN0836	carbonate dehydrat
19	359.5	14.8	260	2	I52951	carbonate dehydrat
20	359.5	14.8	261	1	CRHU1	carbonate dehydrat
21	359.5	14.8	261	2	JN0835	carbonate dehydrat
22	354	14.6	260	1	CRH01D	carbonate dehydrat
23	354	14.5	260	2	A43641	carbonate dehydrat
24	347	14.3	309	2	I51900	carbonic anhydrase
25	343.5	14.2	260	1	CRHU3	carbonate dehydrat
26	342	14.1	304	2	I59261	carbonate dehydrat
27	338.5	14.0	259	2	A22612	carbonate dehydrat
28	333	13.7	334	2	T16772	hypothetical prote
29	328.5	13.6	305	1	CRHU5	carbonate dehydrat

Db 421 GLLFAVTSVAFLVQMRQRRGKGGVSYRPAEVAETGA 459
|||||
RESULT 2
A29993
carbonate dehydratase (EC 4.2.1.1) VI - sheep
C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
C:Date: 15-Dec-1988 #sequence_revision 15-Dec-1988 #text_change 09-Jul-2004
C:Accession: A29993
R:Fernley, R.T.; Wright, R.D.; Coghlan, J.P.
Biochemistry 27, 2815-2820, 1988
A:Title: Complete amino acid sequence of ovine salivary carbonic anhydrase.
A:Reference number: A29993; PMID:88294021; PMID:3135834
A:Accession: A29993
A:Molecule type: protein
A:Residues: 1-307 <PER>
A:Cross-references: UNIPROT:P08060; UNIPARC:UPI0000017604D
A:Experimental source: parotid gland
A:Note: 63-Met and 297-Ile were also found
C:Superfamily: carbonate dehydratase; carbonic anhydrase homology
C:Keywords: carbon-oxygen lyase; hydro-lyase; zinc
F:6-261/Domain: carbonic anhydrase homology <CAH>
Query Match 21.7%; Score 526; DB 2; Length 307;
Best Local Similarity 43.2%; Pred. No. 2.5e-26; Indels 8; Gaps 4;
Matches 111; Conservative 40; Mismatches 98
QY 141 WRVG----GDPPRVS PACAGRFQSPVDIRPQLAAFCPALRPLELLGFOLPPLPELR 196
Db 6 WYSEGMDAEAWPLEYPKCGRRQSPIDLOMKKVQYNPSLRALNLTGGLWH-GEFPT 64
QY 197 NNGHSVQLTPPGLEMALGPGREYRALQLHLHWG--PGSEHTVEGHRPPAEIHVVH 254
Db 65 NNGHTVQISLPSMTMTSDGTQYLAKQMHFWGSGASSEHTVDCGRVYIEIHVVH 124
QY 255 LSTAFARVDEALGRPGGLAVLAFLF-EGPEENSAYEQLLSRLERAEESGSETQVPGLDI 313
Db 125 YNSKYNSEYEAQKEPDGLAVLAALVEVKDYATENAYYSKFIHLEDIRYAGQSTVLRGLDI 184
QY 314 SALLPSDFSRFYQYEGSLTPPCACQGVITVFNQTVMLSAKOLHTLSDTLWGPGDSRLQL 373
Db 185 EDMPLGDLRYYSYLGSLTPPCCTENVHFWVADTVKLSKTQVEKLENSLLHNQNTION 244
QY 374 NFRATQPLNGRVIEASF 390
Db 245 DYRRTQPLNHRVVEANF 261
RESULT 3
S71877
carbonate dehydratase (EC 4.2.1.1) isoform VI precursor - bovine
N:Alternate names: carbonic anhydrase VI
C:Species: Bos primigenius taurus (cattle)
C:Date: 23-Jan-1998 #sequence_revision 06-Feb-1998 #text_change 09-Jul-2004
C:Accession: S71877; S03862
R:Jiang, W.; Woitach, J.T.; Gupta, D.
Biochem. J. 318, 291-296, 1996
A:Title: Sequence of bovine carbonic anhydrase VI: potential recognition sites for N-acetylglucosaminyl transferase.
A:Reference number: S71877; PMID:96358528; PMID:8761494
A:Accession: S71877
A:Molecule type: mRNA
A:Residues: 1-319 <JIA>
A:Cross-references: UNIPROT:P18915; UNIPARC:UPI00000126DF5; EMBL:X96503; NID:g1526571; PI
A:Experimental source: submaxillary gland
R:Fernley, R.T.; Darling, P.; Aldred, P.; Wright, R.D.; Coghlan, J.P.
Biochem. J. 259, 91-96, 1989
A:Title: Tissue and species distribution of the secreted carbonic anhydrase isoenzyme.
A:Reference number: S03862; PMID:89246331; PMID:2497732
A:Accession: S03862
A:Molecule type: protein
A:Residues: 15,'S',17-36,'X',38-39 <PER>
A:Cross-references: UNIPARC:UPI000017604C

C:Genetics:
A:Gene: CAH6
C:Function:
A:Description: catalyzes the reversible dissociation of carbonic acid to carbon dioxide and water
C:Superfamily: carbonate dehydratase; carbonic anhydrase homology
C:Keywords: carbon-oxygen lyase; glycoprotein; hydro-lyase; metalloprotein; zinc
F:1-14/Domain: signal sequence #status predicted <SIG>
F:15-319/Product: carbonic anhydrase isoform VI #status experimental <MAT>
F:18-273/Domain: carbonic anhydrase homology <CAH>
F:37-219/Disulfide bonds: #status predicted
F:62,251/Binding site: carboxylate (Asn) (covalent) #status predicted
F:106,108,133/Binding site: zinc (His) #status predicted
Query Match 21.1%; Score 511; DB 2; Length 319;
Best Local Similarity 42.0%; Pred. No. 2.4e-25;
Matches 108; Conservative 40; Mismatches 97; Indels 12; Gaps 4;
QY 137 DOSHWRYGGDPWPVRVSPACAGRFQSPVDIRPQLAAFCPALRPLELLGFOLPPLPELR 196
Db 26 DEKHWRLE-----QYPCGGTRQSPIDLOMKKVRYNPSLRALNLTGGLRQ-GEFPT 76
QY 197 NNGHSVQLTPPGLEMALGPGREYRALQLHLHWG--AAGRPGSEHTVEGHRPPAEIHVVH 254
Db 77 NNGHTVQISLPSMRMTTSDGTQYLAKQMHFWGSGDSSEHTVDCGRVYIEIHVVH 136
QY 255 LSTAFARVDEALGRPGGLAVLAFLF-EGPEENSAYEQLLSRLERAEESGSETQVPGLDI 313
Db 137 YHSKYGSVEAQNEPDGLAVLAALVEVKDYATENAYYSKFIHLEDIRYAGQSTVLRDLDI 196
QY 314 SALLPSDFSRFYQYEGSLTPPCACQGVITVFNQTVMLSAKOLHTLSDTLWGPGDSRLQL 373
Db 197 QMPLGDLRYYSYLGSLTPPCCTENVHFWVADTVKLSKTQVEKLENSLLHNQNTION 256
QY 374 NFRATQPLNGRVIEASF 390
Db 257 NYRSTQPLNHRVVEANF 273
RESULT 4
CRHU6
carbonate dehydratase (EC 4.2.1.1) VI precursor - human
N:Alternate names: carbonic anhydrase VI; salivary carbonic anhydrase; secreted carbonic
C:Species: Homo sapiens (man)
C:Date: 31-May-1991 #sequence_revision 05-May-1995 #text_change 02-Sep-1997
C:Accession: A37917
R:Aldred, P.; Fu, P.; Barrett, G.; Penschow, J.D.; Wright, R.D.; Coghlan, J.P.; Fernley,
Biochemistry 30, 569-575, 1991
A:Title: Human secreted carbonic anhydrase: cDNA cloning, nucleotide sequence, and hybrid
A:Reference number: A37917; PMID:91105141; PMID:1899030
A:Accession: A37917
A:Molecule type: mRNA
A:Residues: 1-308 <ALD>
A:Cross-references: UNIPARC:UPI0000172F8C; GB:M57892; GB:J05305
A:Experimental source: salivary gland
A:Note: the authors translated the codon GAG for residue 248 as Gln
C:Genetics:
A:Gene: GDB:CA6
A:Cross-references: GDB:125350; OMIM:114780
A:Map position: 1p36-1p36
C:Function:
A:Description: catalyzes the reversible dissociation of carbonic acid to carbon dioxide and
A:Note: this form is expressed in salivary, parotid, and submandibular glands; it is ex
C:Superfamily: carbonate dehydratase; carbonic anhydrase homology
C:Keywords: carbon-oxygen lyase; glycoprotein; hydro-lyase; metalloprotein; pyroglutamic
F:1-17/Domain: signal sequence #status predicted <SIG>
F:18-308/Product: carbonic anhydrase VI #status predicted <MAT>
F:23-278/Domain: carbonic anhydrase homology <CAH>
F:18/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predicted
F:42-224/Disulfide bonds: #status predicted
F:67,256/Binding site: carboxylate (Asn) (covalent) #status predicted
F:111,113,138/Binding site: zinc (His) #status predicted
Query Match 20.3%; Score 492; DB 1; Length 308;

```

Db      173  G-KEHVLKAVTEILQIQYKGSKTIPCFNENTNTLLPDLRLRWVYEGSLTIPPCSEGV
Qy      341  IWTVFNQTVMLSAKOL-----HTLSDTLWGPCDSRLQLNFRATQPLNGRVIEASF 3
Db      232  TWILFRYPLTISOLQIEEFRRLRTHVKGAEIVGCGDGLGDNFRPTQPLSDRVIRAAF 20
RESULT 7
CRHU7
Carbonate dehydratase (EC 4.2.1.1) VII - human
N:Alternate names: carbonic anhydrase VII
C:Species: Homo sapiens (man)
C:Date: 10-Feb-1995 #sequence_revision 05-May-1995 #text_change 09-Jul-2000
C:Accession: A55272
R:Montgomery, J.C.; Venta, P.J.; Eddy, R.L.; Fukushima, Y.S.; Shows, T.B.;
Genomics 11, 835-848, 1991
A:Title: Characterization of the human gene for a newly discovered carbonic
A:Reference number: A55272; MUID:92147127; PMID:1783392
A:Accession: A55272
A:Molecule type: DNA
A:Residues: 1-264 <MON>
A:Cross-references: UNIPROT:P43166; UNIPARC:UPI000000D814; GB:M76423; NID:
A:Note: sequence extracted from NCBI backbone (NCBIN:80199, NCBIN:80201, N
C:Genetics:
A:Gene: GDB:CA7
A:Cross-references: GDB:119741; OMIM:114770
A:Map position: 16q22.1-16q22.1
A:Introns: 14/1; 80/1; 119/3; 151/3; 172/3; 224/3
A:Function:
A:Description: catalyzes the reversible dissociation of carbonic acid to c

```


A;Title: Comparison of the 5' regions of human and mouse carbonic anhydrase II genes and
 A;Reference number: A90655; MUID:86077780; PMID:3000449
 A;Accession: B23202
 A;Molecule type: DNA
 A;Residues: 1-77 <VE2>
 A;Cross-references: UNIPARC:UPI0000172F83
 A;Note: the authors translated the codon CAG for residue 39 as His
 R;Curtis, P.J.; Withers, E.; Demuth, D.; Watt, R.; Venta, P.J.; Tashian, R.E.
 Gene 25, 325-332, 1983
 A;Title: The nucleotide sequence and derived amino acid sequence of cDNA coding for mouse
 A;Reference number: A01143; MUID:84109569; PMID:6420240
 A;Accession: A01143
 A;Molecule type: mRNA
 A;Residues: 2-38, 'H', 40-260 <CUR>
 A;Cross-references: UNIPARC:UPI0000029EE8; GB:K00811; GB:K00812; GB:M11830; NID:gl92333;
 A;Note: Initiator Met not shown
 R;Curtis, P.J.
 J. Biol. Chem. 258, 4459-4463, 1983
 A;Title: Cloning of mouse carbonic anhydrase mRNA and its induction in mouse erythrocyte
 A;Reference number: A20539; MUID:83161023; PMID:6187736
 A;Accession: A20539
 A;Molecule type: mRNA
 A;Residues: 155-178;214-240 <CU2>
 A;Cross-references: UNIPARC:UPI0000172F84; UNIPARC:UPI0000172F85
 R;Venta, P.J.; Montgomery, J.C.; Wiebauer, K.; Hewett-Emmett, D.; Tashian, R.E.
 Ann. N. Y. Acad. Sci. 429, 309-323, 1984
 A;Title: Organization of the mouse and human carbonic anhydrase II genes.
 A;Reference number: I51949; MUID:84255152; PMID:6331255
 A;Accession: I51949
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: mRNA
 A;Residues: 241-260 <RES>
 A;Cross-references: UNIPARC:UPI000000045F; GB:M25944; NID:gl99078; PIDN:AAA39505.1; PID:
 C;Genetics:
 A;Gene: Cat-2
 A;Introns: 12/1; 78/1; 117/3; 144/1; 169/3; 221/3
 C;Superfamily: carbonate dehydratase; carbonic anhydrase homology
 C;Keywords: carbon-oxygen lyase; hydro-lyase; zinc
 F;5-259/Domain: carbonic anhydrase homology <CAH>
 F;94,96,119/Binding site: zinc (His) #status predicted

Query Match 15.7%; Score 380.5; DB 1; Length 260;
 Best Local Similarity 35.6%; Pred. No. 3.5e-17;
 Matches 93; Conservative 41; Mismatches 112; Indels 15; Gaps 8;

Qy 140 HWRY---GGDPWPVRVSPACAGRFQSPVDIRPQLAAFCPALRPLUELGFLQPLPELRLR 196
 ||||| : : : : : ||||| : : : : : ||||| : : : : :
 Db 4 HWGSKINGPENWHKDFPIANGDRQSPVDITATAQHPALQPL-LISYD--KAASKSIV 60
 ||||| : : : : : ||||| : : : : : ||||| : : : : :
 Qy 197 NNGHSVQLTLPPLGLEMAL---GP-GREYRALQLHLHWAAGRPQSGSEHTVEGHRFPALIHV 252
 ||||| : : : : : ||||| : : : : : ||||| : : : : :
 Db 61 NNGHSFNVEFDDSQDQNAVLKGGPLSDSYRLIQFHFHMGSSDQSGSEHTVKNKKYAAELHL 120
 ||||| : : : : : ||||| : : : : : ||||| : : : : :
 Qy 253 VHLSTAFARVDEALGRPGGLAVLAFLKEEGPEENSAYEOLLSRLLEAESEGTQVPGLD 312
 ||||| : : : : : ||||| : : : : : ||||| : : : : :
 Db 121 VHWNTKYGDFGKAVQCPDGLAVGLFKIGP-ASQGLQKVLEALHSITKKGKRAAFANFD 179
 ||||| : : : : : ||||| : : : : : ||||| : : : : :
 Qy 313 ISALLPSDFSRYFOYEGSLTTPPCACQGVIVTFVFNQVMSAKQL---HTLSDTLWPGDS 369
 ||||| : : : : : ||||| : : : : : ||||| : : : : :
 Db 180 PCSLLPLGNDL-YWYTPGSLTTPPLELCEVTWIVREPITVSSEQMSHFRTLNFNEEGDAEE 238
 ||||| : : : : : ||||| : : : : : ||||| : : : : :
 Qy 370 RLQINFRATQPLNGRVITASF 390
 : : : : : ||||| : : : : :
 Db 239 AMVDNWRPAQLKVRKIKASF 259
 : : : : : ||||| : : : : :
 RESULT 11
 CRW02
 N;Alternate names: carbonic anhydrase II; hepatic carbonic anhydrase
 C;Species: Homo sapiens (man)
 C;Date: 07-May-1981 #sequence revision 05-May-1995 #text change 09-Jul-2004
 C;Accession: A27175; A23202; A92194; A92147; I37214; I51863; I51871; A01141

R;Murakami, H.; Marelich, G.P.; Grubb, J.H.; Kyle, J.W.; Sly, W.S.
 Genomics 1, 159-166, 1987
 A;Title: Cloning, expression, and sequence homologies of cDNA for human carbonic anhydrase
 A;Reference number: A27175; MUID:88085190; PMID:3121496
 A;Accession: A27175
 A;Molecule type: mRNA
 A;Residues: 1-260 <MUR>
 A;Cross-references: UNIPROT:P00918; UNIPARC:UPI0000110BA2; GB:J03037; NID:gl79771; PIDN:
 R;Venta, P.J.; Montgomery, C.; Hewett-Emmett, D.; Tashian, R.E.
 Biochim. Biophys. Acta 826, 195-201, 1985
 A;Title: Comparison of the 5' regions of human and mouse carbonic anhydrase II genes and
 A;Reference number: A90655; MUID:86077780; PMID:3000449
 A;Accession: A23202
 A;Molecule type: DNA
 A;Residues: 1-77 <VEN>
 A;Cross-references: UNIPARC:UPI000016A650; GB:X03251; GB:M18100; GB:M77181; NID:gl79778;
 R;Henderson, L.E.; Henriksson, D.; Nyman, P.O.
 J. Biol. Chem. 251, 5457-5463, 1976
 A;Title: The primary structure of human carbonic anhydrase C.
 A;Reference number: A92194; MUID:77006079; PMID:823150
 A;Accession: A92194
 A;Molecule type: protein
 A;Residues: 2-260 <HEN>
 A;Cross-references: UNIPARC:UPI0000110396
 R;Lin, K.T.D.; Deutsch, H.F.
 J. Biol. Chem. 249, 2329-2337, 1974
 A;Title: Human carbonic anhydrases. XII. The complete primary structure of the C isozyme
 A;Reference number: A92147; MUID:74143468; PMID:4207120
 A;Accession: A92147
 A;Molecule type: protein
 A;Residues: 2-260 <LIN>
 A;Cross-references: UNIPARC:UPI0000110396
 R;Montgomery, J.C.; Venta, P.J.; Tashian, R.E.; Hewett-Emmett, D.
 Nucleic Acids Res. 15, 4687, 1987
 A;Title: Nucleotide sequence of human liver carbonic anhydrase II cDNA.
 A;Reference number: I37214; MUID:87231043; PMID:3108857
 A;Accession: I37214
 A;Status: translated from GB/EMBL/DBJ
 A;Molecule type: mRNA
 A;Residues: 1-260 <RE3>
 A;Cross-references: UNIPARC:UPI0000110BA2; EMBL:Y00339; NID:g29586; PIDN:CAA68426.1; PID:
 A;Experimental source: liver
 A;Note: submitted to the EMBL/GenBank/DBJ databases by David Hewett-Emmett 01-JUL-1987
 R;Venta, P.J.; Welty, R.J.; Johnson, T.M.; Sly, W.S.; Tashian, R.E.
 Am. J. Hum. Genet. 49, 1082-1090, 1991
 A;Title: Carbonic anhydrase II deficiency syndrome in a Belgian family is caused by a po
 e.
 A;Reference number: I51863; MUID:92026087; PMID:1928091
 A;Accession: I51863
 A;Status: translation not shown; translated from GB/EMBL/DBJ
 A;Molecule type: DNA
 A;Residues: 1-260 <RES>
 A;Cross-references: UNIPARC:UPI0000110BA2; GB:M77181; NID:gl79778; PIDN:AAA51909.1; PID:
 A;Note: the complete nucleotide sequence is not shown
 R;Hu, P.Y.; Ernst, A.R.; Sly, W.S.; Venta, P.J.; Skaggs, L.A.; Tashian, R.E.
 Am. J. Hum. Genet. 54, 602-608, 1994
 A;Title: Carbonic anhydrase II deficiency: single-base deletion in exon 7 is the predomi
 A;Reference number: I51871; MUID:94175074; PMID:8128957
 A;Accession: I51871
 A;Status: translated from GB/EMBL/DBJ
 A;Molecule type: mRNA
 A;Residues: 225-242 <RE2>
 A;Cross-references: UNIPARC:UPI00000004F0; GB:S69526; NID:g545850; PIDN:AAB30170.1; PID:
 A;Note: wild type shown; mutant contains frameshift after residue 226
 R;Eriksson, A.E.; Jones, T.A.; Liljas, A.
 submitted to the Brookhaven Protein Data Bank, February 1989
 A;Reference number: A50085; PDB:1CA2
 A;Contents: annotation; X-ray crystallography, 2.0 angstroms, residues 4-259
 R;Liljas, A.; Kannan, K.K.; Bergsten, P.C.; Waara, I.; Fridborg, K.; Strandberg, B.; Car
 Nature New Biol. 235, 131-137, 1972
 A;Title: Crystal structure of human carbonic anhydrase C.
 A;Reference number: A93404; MUID:72111787; PMID:4621826
 A;Contents: annotation; X-ray crystallography, 2.0 angstroms

RESULT 14

[illegible]

Search completed: February 12, 2006, 19:56:28
Job time : 45 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 12, 2006, 19:48:36 ; Search time 247 Seconds

(without alignments)
1311.083 Million cell updates/sec

Title: US-09-967-237B-2

Perfect score: 2424

Sequence: 1 MAPLCPSPWLLIPAPAG.....RRGTGGSVYRPAEVAETGA 459

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Uniprot_05.80.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2424	100.0	459	1	CAH9_HUMAN
2	2420	99.8	459	2	Q5T4R1_HUMAN
3	1596	65.8	437	1	CAH9_MOUSE
4	562	23.2	337	1	CAH14_HUMAN
5	562	23.2	337	2	Q5TB24_HUMAN
6	558	23.0	337	1	CAH14_MOUSE
7	537.5	22.2	354	1	CAH12_HUMAN
8	526	21.7	307	1	CAH6_SHEEP
9	525	21.7	343	2	Q53YB5_HUMAN
10	519	21.4	525	2	Q6GQ30_XENLA
11	515	21.2	320	1	CAH6_CAFPA
12	512	21.1	355	1	CAH12_RABIT
13	512	21.1	526	2	Q4SZWI_TETNG
14	511	21.1	319	1	CAH6_BOVIN
15	507	20.9	282	2	Q4RVR1_TETNG
16	507	20.9	313	2	Q5FC00_HUMAN
17	498	20.5	323	2	Q58HA5_RAT
18	496	20.5	354	1	CAH12_MOUSE
19	493	20.3	344	2	Q8K2J1_MOUSE
20	488	20.1	308	1	CAH6_HUMAN
21	467.5	19.3	317	1	CAH6_MOUSE
22	465.5	19.2	317	2	Q7TNG9_MOUSE
23	465.5	19.2	325	2	Q80YB7_MOUSE
24	452	18.6	305	2	Q4RR10_TETNG
25	448.5	18.5	229	2	Q4SL23_TETNG
26	446.5	18.4	314	2	Q6DBS1_BRARE
27	445.5	18.4	359	2	Q4SNCO_TETNG
28	437.5	18.0	302	2	Q4FCT9_SQUAC
29	435.5	18.0	268	2	Q4RRY2_TETNG
30	430.5	17.8	312	2	Q7SYW3_XENLA
31	427.5	17.6	260	2	Q6PFU7_BRARE

32	426.5	17.6	258	2	QSMCNO_PSEAM
33	425.5	17.6	297	2	Q6R4A1_ONCMY
34	422.5	17.4	258	2	Q6JRS3_OREMO
35	422.5	17.4	224	1	CAH15_MOUSE
36	422	17.4	259	1	CAH2_TRIHK
37	419.5	17.3	290	1	CAH8_MOUSE
38	419	17.3	309	2	Q6NWD3_BRARE
39	412	17.0	289	1	CAH8_HUMAN
40	409	16.9	264	1	CAH7_HUMAN
41	409	16.9	264	1	CAH7_MOUSE
42	409	16.9	264	2	Q541F0_HUMAN
43	409	16.9	290	2	Q5PPN4_RAT
44	402.5	16.6	260	1	CAH2_BRARE
45	401	16.5	260	2	Q68YC2_ONCMY

ALIGNMENTS

RESULT 1

ID	CAH9_HUMAN	STANDARD;	PRT;	459 AA.
AC	Q16790;			
DT	15-JUL-1999 (Rel. 38, Last sequence update)			
DT	15-JUL-1999 (Rel. 38, Last sequence update)			
DT	13-SEP-2005 (Rel. 48, Last annotation update)			
DE	Carbonic anhydrase IX precursor (EC 4.2.1.1) (Carbonate dehydratase IX) (CA-IX) (CAIX) (Membrane antigen MN) (P54/58N) (Renal cell carcinoma-associated antigen G250) (RCC-associated antigen G250) (pMW1).			
DE	carcinoma-associated antigen G250) (RCC-associated antigen G250) (pMW1).			
GN	Name=CA9; Synonyms=G250, MN;			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;			
OC	Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	NUCLEOTIDE SEQUENCE, AND CHARACTERIZATION.			
RC	TISSUE=Carcinoma;			
RX	MEDLINE=94366734; PubMed=8084592;			
RA	Pastorek J., Pastorekova S., Callebaut I., Mornon J.-P., Zelnik V., Opavsky R., Zato Ovicova M., Liao S., Portetelle D., Stanbridge E.J., Zavada J., Burny A., Kettmann R.;			
RA	"Cloning and characterization of MN, a human tumor-associated protein with a domain homologous to carbonic anhydrase and a putative helix-loop-helix DNA binding segment.";			
RL	Oncogene 9:2877-2888(1994).			
RN	[2]			
RP	NUCLEOTIDE SEQUENCE, AND VARIANT VAL-33.			
RC	TISSUE=Renal cell carcinoma;			
RX	MEDLINE=20175484; PubMed=10709109;			
RA	Grabmaier K., Vissers J.L.M., De Weijert M.C.A., Brakenhoff R.H., Noessner E., Oosterwijk-Wakka J.C., Van Bokhoven A., Brakenhoff R.H., Noessner E., Mulders P.A., Merks J.G., Figdor C.G., Adema G.J., Oosterwijk E.;			
RA	"Molecular cloning and immunogenicity of renal cell carcinoma-associated antigen G250.";			
RT	Int. J. Cancer 85:865-870(2000).			
RL	[3]			
RP	NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].			
RC	TISSUE=Colon;			
RX	MEDLINE=12477932; PubMed=1073/pnas.242603899;			
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.B., Brownstein M.J., Uadin T.B., Toshiyuki S., Carninci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullen S.H., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,			

DR GO; GO:0004089; F:carbonate dehydratase activity; IEA.
 DR GO; GO:0008270; F:zinc ion binding; IEA.
 DR GO; GO:0006730; P:one-carbon compound metabolism; IEA.
 SQ SEQUENCE 459 AA; 49697 MW; BA67195483F0F5CE CRC64;

Query Match 99.8%; Score 2420; DB 2; Length 459;
 Best Local Similarity 99.8%; Pred. No. 2.7e-134;
 Matches 458; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAPLCPSPFWLPLLIAPAPAGLTGVLQALLSLLLMPVHPQRLPRMQEDSPPLGGSGSGDDPL 60
 DB 1 MAPLCPSPFWLPLLIAPAPAGLTGVLQALLSLLLMPVHPQRLPRMQEDSPPLGGSGSGDDPL 60

QY 61 GEEDLPSEEDSPREDPPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLP 120
 DB 61 GEEDLPSEEDSPREDPPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLP 120

QY 121 DPQFPQNNAHKDKGDDQSHWRYGDPWPVRSVSPACAGRFQSPVDIRPQLAAFCPALRPL 180
 DB 121 DPQFPQNNAHKDKGDDQSHWRYGDPWPVRSVSPACAGRFQSPVDIRPQLAAFCPALRPL 180

QY 181 ELLGFQPLPPLPELRLNNGHVSQVLTLPGLBMAALPGREYRALQLHLHWGAAGRPGSEHT 240
 DB 181 ELLGFQPLPPLPELRLNNGHVSQVLTLPGLBMAALPGREYRALQLHLHWGAAGRPGSEHT 240

QY 241 VEGHFPFAEIHVHLSVAFARVDEALGRPGGLAVLAFLPEGPEENAYEQLLSLRLBEIA 300
 DB 241 VEGHFPFAEIHVHLSVAFARVDEALGRPGGLAVLAFLPEGPEENAYEQLLSLRLBEIA 300

QY 301 EEGSETQVGLDLSALLPDSFSRYFQVYEGSLTTPPCAGVITWTFNQTVMLSAKQLHTLS 360
 DB 301 EEGSETQVGLDLSALLPDSFSRYFQVYEGSLTTPPCAGVITWTFNQTVMLSAKQLHTLS 360

QY 361 DTLGPGDSRLQLNFRATQPLNGRVIEASFAGVDSSPRAAEPVQLNSCLAAGDILALVF 420
 DB 361 DTLGPGDSRLQLNFRATQPLNGRVIEASFAGVDSSPRAAEPVQLNSCLAAGDILALVF 420

QY 421 GLLFAVTSVAFVQMRQRHRTGKGVSRYRPAEVAETGA 459
 DB 421 GLLFAVTSVAFVQMRQRHRTGKGVSRYRPAEVAETGA 459

RESULT 3

CAH9 MOUSE STANDARD; PRT; 437 AA.

AC Q8VHB5; Q8K1G1; Q8VDE4;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 13-SEP-2005 (Rel. 48, Last annotation update)
 DE Carbonic anhydrase IX precursor (EC 4.2.1.1) (Carbonate dehydratase IX) (CA-IX) (CAIX) (Membrane antigen MN homolog).
 GN Name=Ca9; Synonyms=Ca9;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
 OC Muridae; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP NUCLEOTIDE SEQUENCE (ISOFORM 1).
 RC STRAIN=129/Ola;
 RX PubMed=14604546; DOI=10.1016/j.jim.2003.08.011;
 RA Zlat'ovicova M., Tarabkova K., Svasatova E., Gibadulinova A., Mucha V.,
 RA Jakubickova L., Biesova Z., Rafajova M., Ortova Gut M.O., Parkkila S.,
 RA Parkkila A.-K., Waheed A., Sly W.S., Horak I., Pastorek J.,
 RA Pastorekova S.;
 RT "Monoclonal antibodies generated in carbonic anhydrase IX-deficient mice recognize different domains of tumour-associated hypoxia-induced carbonic anhydrase IX";
 RL J. Immunol. Methods 282:117-134(2003).
 RN [2]
 RP NUCLEOTIDE SEQUENCE (ISOFORM 1).
 RA Ortova M.;
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.

[3]
 RN NUCLEOTIDE SEQUENCE (ISOFORM 2).
 RP STRAIN=ICR; TISSUE=Small intestine;
 RA Wang Y.P., Yoshikawa K., Kozaki K., Miyaishi O., Nakagawa A.,
 RA Muramatsu H., Kawada Y., Uchida K., Nishikawa N., Saga S.;
 RT "Alternative spliced mRNA coding for MN/CA9";
 RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: Reversible hydration of carbon dioxide (By similarity).
 CC -1- CATALYTIC ACTIVITY: H(2)CO(3) = CO(2) + H(2)O.
 CC -1- COFACTOR: Zinc (By similarity).
 CC -1- SUBUNIT: Forms oligomers linked by disulfide bonds (By similarity).
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=1;
 CC IsoId=Q8VHB5-1; Sequence=Displayed;
 CC Name=2;
 CC IsoId=Q8VHB5-2; Sequence=VSP_007409, VSP_007410;
 CC -1- SIMILARITY: Belongs to the eukaryotic-type carbonic anhydrase family.
 CC -----
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use as long as its content is in no way modified and this statement is not removed.
 CC -----
 DR EMBL; AY049077; AAL14193.1; -; Genomic_DNA.
 DR EMBL; AJ245857; CAC80975.1; -; mRNA.
 DR EMBL; AB086322; BAC00816.1; -; mRNA.
 DR HSP; O43570; 1JDO.
 DR Ensembl; ENSMUSG0000028463; Mus musculus.
 DR MGI; MGI-2447188; Car9.
 DR GO; GO:0005615; C:extracellular space; TAS.
 DR GO; GO:0016021; C:integral to membrane; TAS.
 DR InterPro; IPR001148; Euk_Coanhd.
 DR PANTHER; PTHR18952; Euk_Coanhd; 1.
 DR Pfam; PF00194; Carb_anhydrase; 1.
 DR PRODOM; PD000865; Euk_Coanhd; 1.
 DR PROSITE; PS00162; Euk_CO2_ANHYDRASE; 1.
 DR KW Alternative splicing; Glycoprotein; Lyase; Metal-binding; Signal; Transmembrane; Zinc.
 FT SIGNAL 1 31 Potential.
 FT CHAIN 32 437 Carbonic anhydrase IX.
 FT TRANSMEM 391 411 Potential.
 FT METAL 205 205 Zinc (catalytic) (By similarity).
 FT METAL 207 207 Zinc (catalytic) (By similarity).
 FT METAL 230 230 Zinc (catalytic) (By similarity).
 FT CARBOHYD 325 325 N-linked (GlcNAc...) (Potential).
 FT VARSPPLIC 282 282 G -> V (in isoform 2).
 FT FTID=VSP_007409.
 FT VARSPPLIC 283 437 Missing (in isoform 2).
 FT FTID=VSP_007410.
 SQ SEQUENCE 437 AA; 47265 MW; 88F23380DCD35344 CRC64;

Query Match 65.8%; Score 1596; DB 1; Length 437;
 Best Local Similarity 63.4%; Pred. No. 7e-86;
 Matches 320; Conservative 36; Mismatches 79; Indels 26; Gaps 6;

QY 1 MAPLCPSPFWLPLLIAPAPAGLTGVLQALLSLLLMPVHPQRLPRMQEDSPPLGGSGSGDDPL 60
 DB 1 MASLCPSPFWLSTPAP-----TAQLLLFLLLQVSAQPGQLSGMQGEPGLSDSSGDE-L 55

QY 61 GEEDLPSEEDSPREDPPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLP 120
 DB 56 GVDVLPSEEDAPPEADP-----PDGPPPEVNSDRMBESLGLEDLSTPEAP- 102

QY 121 DPQFPQNNAHKDKGDDQSHWRYGDPWPVRSVSPACAGRFQSPVDIRPQLAAFCPALRPL 180
 DB 103 ---EHSQSGSHGDERGGHSHWSYGGTLLWPQVSPACAGRFQSPVDIRLERTAFCTLQPL 159

QY 181 ELLGFQPLPPLPELRLNNGHVSQVLTLPGLBMAALPGREYRALQLHLHWGAAGRPGSEHT 240
 DB 181 ELLGFQPLPPLPELRLNNGHVSQVLTLPGLBMAALPGREYRALQLHLHWGAAGRPGSEHT 240

Db 160 ELLGVELQPLPELSLNNGHTVQLTLPGLKVALGPGQRYRALQLHLHWGTSDDPGSEHT 219

QY 241 VEGHREPAEIHVVHLSTAFARVDEALPGGLAVLAFLAEQPEENSAYEQLLSLREITA 300

Db 220 VNGHREPAEIHVVHLSTAFSELHEALPGGLAVLAFLAEQPEENSAYEQLLSHLEITS 279

QY 301 EGSETOVGLDISALLPDSFRYFOYEGSLTTPPCACQGVITVFNQTVMSAKQLHTLS 360

Db 280 EBGSEKIEPLGVDVALLPDSLSRYRYEGSLTTPPCQGVITVFNQTVMSAKQLHTLS 339

QY 361 DTLWGPGRDLQNLFRATOPNGRVTEAGFPAGVDSSPRAAPVOLNSCLAAGDIIALVF 420

Db 340 VSLWGPGRDLQNLFRATOPNGRVTEAGFPRAEDSSP---FPVHNSCFTAGDIIALVF 396

QY 421 GLLFAVTSVAFVLQVRRQHR--GTKGGVGYRPAEVAETGA 459

Db 397 GLLFAVTSIAFLQLRQRHRSRGTCDRVSYSPAEMTETGA 437

RESULT 4

CAH14 HUMAN

ID CAH14 HUMAN STANDARD; PRT; 337 AA.

AC Q9ULX7; O8NCF4;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 13-SEP-2005 (Rel. 48, Last annotation update)

DE Carbonic anhydrase XIV precursor (SC 4.2.1.1) (Carbonate dehydratase XIV) (CA-XIV)

GN Name=CA14; ORFNames=UNQ690/PRO1335;

OS Homo sapiens (Human)

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;

OC Homo.

OX NCBI_TaxID=9606;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RX MEDLINE=99443874; PubMed=10512682; DOI=10.1006/geno.1999.5938; Fujikawa-Adachi K., Nishimori I., Taguchi T., Onishi S.; "Human carbonic anhydrase XIV (CA14): cDNA cloning, mRNA expression, and mapping to chromosome 1.," Genomics 61:74-81(1999).

RN [2]

RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].

RX MEDLINE=22987296; PubMed=12975309; DOI=10.1101/gr.1293003; Clark H.F., Gurney A.L., Abaya E., Baker K., Baldwin D.T., Brush J., Chen J., Chow B., Chui C., Crowley C., Currell B., Deuel B., Dowd P., Eaton D., Foster J.S., Grimaldi C., Gu Q., Hass P.E., Heldens S., Huang A., Kim H.S., Klimowski L., Jin Y., Johnson S., Lee J., Lewis L., Liao D., Mark M.R., Robbie E., Sanchez C., Schoenfeld J., Seshagiri S., Simmons L., Singh J., Smith V., Stinson J., Vagts A., Vandlen R.L., Watanabe C., Wieand D., Woods K., Xie M.-H., Yansura D.G., Yi S., Yu G., Yuan J., Zhang M., Zhang Z., Goddard A.D., Wood W.I., Godowski P.J., Gray A.M.; "The secreted protein discovery initiative (SPDI), a large-scale effort to identify novel human secreted and transmembrane proteins: a bioinformatics assessment.," Genome Res. 13:2265-2270(2003).

RN [3]

RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].

RX PubMed=14702039; DOI=10.1038/ng1285; Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R., Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H., Sekine M., Obayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S., Yamamoto J.-I., Saito K., Kawai Y., Isono Y., Nakamura Y., Nagahara K., Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M., Shiratori A., Sudo H., Hosoi T., Kaku Y., Kodaira H., Kondo H., Sugawara M., Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E., Omura Y., Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M., Yamazaki M., Ninomiya K., Ishibashi T., Yamashita H., Murakawa K., Fujimori K., Tanai H., Kimata S., Watanabe S., Hirakawa S., Chiba Y., Ishida S., Ono Y., Takiguchi S., Watanabe S., Yosida M., Hotuta T., Kusano J., Kanehori K., Takahashi-Fujii A., Hara H., Tanase T.-O., Nomura Y., Togliya S., Komai F., Hara R., Takeuchi K., Arita M.,

Imose N., Muesashino K., Yuuki H., Oshima A., Sasaki N., Aotsuka S., Yoshikawa Y., Matsunawa H., Ichihara T., Shiohata N., Sano S., Moriya S., Momiyama H., Satoh N., Takami S., Terashima Y., Suzuki O., Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H., Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B., Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y., Fujimori Y., Komiyama M., Tashiro K., Tanigami A., Fujiwara T., Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y., Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S., Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T., Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Komatsu T., Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T., Takahashi K., Suganuma J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K., Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R., Nakai K., Yada T., Nakamura Y., Ohara O., Isogai H., Sugano S.; "Complete sequencing and characterization of 21,243 full-length human cDNAs.," Nat. Genet. 36:40-45(2004).

RN [4]

RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].

RX Kalline N., Chen X., Rolfs A., Halleck A., Hines L., Eisenstein S., Koundinya M., Raphael J., Moreira D., Kelley T., LaBaer J., Lin Y., Phelan M., Farmer A.; "Cloning of human full-length cDNAs in BD Creator(TM) system donor vector.," Submitted (MAY-2003) to the EMBL/GenBank/DDAJ databases.

RN [5]

RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].

RX TISSUE=Brain, Lung, and Testis; MEDLINE=23388257; PubMed=12477932; DOI=10.1073/pnas.242603899; Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., Altshul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hong L., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz T.E., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullighy S.J., Bosak S.A., McGowan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Gibbs R.A., Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E., Scherch A., Schein J.E., Jones S.J.M., Marra M.A.; "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences.," Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

RN [6]

RP FUNCTION: Reversible hydration of carbon dioxide.

RX CATALYTIC ACTIVITY: H(2)CO(3) = CO(2) + H(2)O.

RN [7]

RP COFACTOR: Zinc (By similarity).

RN [8]

RP SUBCELLULAR LOCATION: Type I membrane protein (Potential).

RN [9]

RP TISSUE SPECIFICITY: High expression in all parts of the central nervous system and lower expression in adult liver, heart, small intestine, colon, kidney, urinary bladder and skeletal muscle.

RN [10]

RP SIMILARITY: Belongs to the eukaryotic-type carbonic anhydrase family.

RN [11]

RP This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use as long as its content is in no way modified and this statement is not removed.

RN [12]

RP EMBL; AB025904; BAA85002.1; -; mRNA.

DR EMBL; AB025904; BAA85002.1; -; mRNA.

DR EMBL; AK074765; BAC1191.1; -; mRNA.

DR EMBL; BT020054; AAV38857.1; -; mRNA.

DR EMBL; BC034412; AAV34412.1; -; mRNA.

DR HSSP; O43570; 1JDO.

DR SMR; Q9ULX7; 20-278.


```

DR DR Ensembl; ENSG00000118298; Homo sapiens.
DR DR HGNC; HGNC:1372; CA14.
DR DR MIM; 604832; -.
DR DR GO; GO:0016021; C: integral to membrane; TAS.
DR DR GO; GO:0004089; F: carbonate dehydratase activity; TAS.
DR DR InterPro; IPR001148; Euk Coanhd.
DR DR PANTHER; PTHR18952; Euk Coanhd; 1.
DR DR Pfam; PF00194; Carb_anhyd; 1.
DR DR ProDom; PD000865; Euk Coanhd; 1.
DR DR PROSITE; PS00162; Euk CO2 ANHYDRASE; FALSE NEG.
KW Glycoprotein; Lyase; Metal-binding; Signal; Transmembrane; Zinc.
FT SIGNAL 1 15 Potential.
FT CHAIN 16 337 Carbonic anhydrase XIV.
FT TOPO_DOM 16 290 Extracellular (Potential).
FT TRANSMEM 291 311 Potential.
FT TOPO_DOM 312 337 Cytoplasmic (Potential).
FT METAL 109 109 Zinc (catalytic) (By similarity).
FT METAL 111 111 Zinc (catalytic) (By similarity).
FT METAL 135 135 Zinc (catalytic) (By similarity).
FT CARBOHYD 213 213 N-linked (GlcNAc...) (Potential).
FT CONFLICT 229 229 V -> A (in Ref. 3).
FT SEQUENCE 337 AA; 37668 MW; 6E101C44EA70A700 CRC64;

Query Match 23.2%; Score 562; DB 1; Length 337;
Best Local Similarity 38.8%; Pred. No. 3.2e-25;
Matches 128; Conservative 47; Mismatches 129; Indels 26; Gaps 7;

QY 136 DDQSHWRY---GGDPWPVRVSPACAGRFQSPVDIRPOLAFCPALRPLELLGFGPLPE 192
DB 17 DGGQHWTYEGPHGDHWPASYPCCGNAQSPIDQTDVTFDPLPALQPHGYPQGPTEP 76
QY 193 LRLRNNGHSVLTLPGLGEMALGPGREYRALQLHLHWAAGRP-GSEHTVEGHRFPPEIH 251
DB 77 LDLENNGHTVQLSLPSTLYLG-GLPRKYAAQLHLHWAAGRP-GSEHTVEGHRFPPEIH 135
QY 252 VVHL-STAFARVDEALGRPGGLAVLAALFLEEGPEENSAVEQLLSRLSEIEAEGSETQVPG 310
DB 136 IVHVDSDYSLSEAAERPGGLAVGLILIEVETKNTAYEHLHLHVRHKKQKTSVPP 195
QY 311 LDISALLPDSFRRYFOYEGSLTTPPCAGQVITVFNQTVMLSAKQLHTSLDSTLWPGDSR 370
DB 196 FNRLLEPKQLGQYFRYNGSLTTPCYQSVLTVFRRSQISMELKQLGTLFSTBEEP 255
QY 371 LQL---NFRATQPLNGRVIEASFPAGVDSPPRAEPVQLNSCLAAGDILALVFLGLFA-- 425
DB 256 SKLLVQVYRALQPLNRMVFASF-----IQAGSYTTGEMLSLGVGLVGLCL 302

QY 426 --VTSVAFVQMRQRHRRGTGGVSYRPAE 453
DB 303 CLLAVYFIARKIRKRLNKRKSVVFTSAQ 332

RESULT 6
CAH14_MOUSE
ID CAH14_MOUSE STANDARD; PRT; 337 AA.
AC Q9WV76;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Carbonic anhydrase XIV precursor (EC 4.2.1.1) (Carbonate dehydratase XIV) (CA-XIV).
GN Mus musculus (Mouse).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Kidney;
RX MEDLINE=92663110; PubMed=10336468; DOI=10.1074/jbc.274.22.15701;
RA Mori K., Ogawa Y., Ebihara K., Tamura N., Tashiro K., Kuwahara T.,
RA Mukoyama M., Sugawara A., Ozaki S., Tanaka I., Nakao K.;
RA "Isolation and characterization of CA XIV, a novel membrane-bound
RA carbonic anhydrase from mouse kidney.";
RA J. Biol. Chem. 274:15701-15705(1999).
RL [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC TISSUE=Retina;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raba S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,

```

RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC -!- FUNCTION: Reversible hydration of carbon dioxide.
 CC -!- CATALYTIC ACTIVITY: H(2)CO(3) = CO(2) + H(2)O.
 CC -!- COFACTOR: Zinc (By similarity).
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein (Potential).
 CC -!- TISSUE SPECIFICITY: Most abundant in the kidney and heart,
 CC followed by the skeletal muscle, brain, lung and liver.
 CC -!- SIMILARITY: Belongs to the eukaryotic-type carbonic anhydrase
 CC family.

 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.

 CC EMBL: AB005450; BAA78709.1; -; mRNA.
 DR EMBL: BC046995; AAH46995.1; -; mRNA.
 DR EMBL: IRJ5; X-ray; A/B-18-278.
 DR PDB: IRJ6; X-ray; A/B-18-278.
 DR Ensembl: ENSMUSG00000038526; Mus musculus.
 DR MGI: MGI:1344341; Carli4.
 DR GO: GO:0005615; C:extracellular space; TAS.
 DR GO: GO:0016021; C:integral to membrane; TAS.
 DR InterPro: IPR001148; Euk_Coanhd.
 DR PANTHER: PTHR18952; Euk_Coanhd; 1.
 DR Pfam: PF00194; Carb_anhydrase; 1.
 DR ProDom: PD000865; Euk_Coanhd; 1.
 DR PROSITE: PS00162; Euk_CO2 ANHYDRASE; FALSE NEG.
 KW 3D-structure; Glycoprotein; Lyase; Metal-binding; Signal;
 KW Transmembrane; Zinc.
 FT SIGNAL 1 15 Potential.
 FT CHAIN 16 337 Carbonic anhydrase XIV.
 FT TOPO_DOM 16 290 Extracellular (Potential).
 FT TRANSMEM 291 311 Potential.
 FT TOPO_DOM 312 337 Cytoplasmic (Potential).
 FT METAL 109 109 Zinc (catalytic) (By similarity).
 FT METAL 111 111 Zinc (catalytic) (By similarity).
 FT METAL 135 135 Zinc (catalytic) (By similarity).
 FT CARBOHYD 213 213 N-linked (GlcNAc...) (Potential).
 SQ SEQUENCE 337 AA; 37505 MW; 32F02F4DB78AC0C9 CRC64;
 Query Match 23.0%; Score 558; DB 1; Length 337;
 Best Local Similarity 39.4%; Pred. No. 5.4e-25;
 Matches 130; Conservative 42; Mismatches 130; Indels 28; Gaps 8;
 QY 136 DQSHWY---GGDPWPVRVSPACGRFOSFVDIRPQLAFCFALRPLLELGFQLPPLPE 192
 DB 17 DGHHTYEGPHGQDHWTPTYPECGDQASPINIQTSDVIFDPLPAVPHGYDQLGTGP 76
 QY 193 LRLRNNGHVSQVLTPLPLEMALGPGREYRALQLHLHWGAAGR-PSGEHTVEGHRFPAETH 251
 DB 77 LDLHNNGHTVQLSLPTLHLG-GLPRKYTAAQLHLHWGQSGLEGSEHQINSEATAAELH 135
 QY 252 VVHL-STAFARVDEALGRGGLAVLAFLFEEGPEENSAYEQLLSRLLEEAEESSETQVPG 310
 DB 136 VVHYSQSYSSSEAAQKQGLAVLGILFVETGENTENPAYDHLISRLHEIRYKDKQTSVPP 195
 QY 311 LDISALLPDSFYSFYEGSLTPPCCAQGVIVTWFVQTVMLSAKQL-----HTLSDTLWGP 366
 DB 196 FSVRELPFPOOLEGFFRYNGLTTPCYQSVLWTFVFNRAQISMGQLEKLTLSSTEEDP 255
 QY 367 GDSRLQLNFRATPLNGRVIEASFPAAGVDSPPRAEPVQVNSCLAAAGDILALVGLLFA- 425
 DB 256 SEPLVQ-NYRVQVPLNQTIFASF-----IQAGPLTYTTGEMGLGVGLIAGC 301
 QY 426 ---VTSVAVLQVWRHRRGRTKGVSYRPA 452

Db 302 LCILLAVYFIAQKIRKRLGNKRSVVFTSA 331
 RESULT 7
 CAH12 HUMAN
 ID CAH12 HUMAN STANDARD; PRT; 354 AA.
 AC Q43570; Q9BWG2;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 13-SEP-2005 (Rel. 48, Last annotation update)
 DE Carbonic anhydrase XII precursor [EC 4.2.1.1] (Carbonate dehydratase
 DE XII) (CA-XII) (Tumor antigen HOM-RCC-3.1.3).
 GN Name=CA12;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
 OC Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP NUCLEOTIDE SEQUENCE (ISOFORM 1), AND CHARACTERIZATION.
 RC TISSUE=Renal cell carcinoma;
 RX MEDLINE=98301622; PubMed=9636197; DOI=10.1073/pnas.95.13.7608;
 RA Tuercio O., Sahin U., Vollmar E., Siemer S., Goettert E., Seitz G.,
 RA Parkkila A.-K., Shah G.N., Grubb J.H., Pfreundschuh M., Sly W.S.;
 RT "Human carbonic anhydrase XII: cDNA cloning, expression, and
 RT chromosomal localization of a carbonic anhydrase gene that is
 RT overexpressed in some renal cell cancers.";
 RL Proc. Natl. Acad. Sci. U.S.A. 95:7608-7613(1998).
 RN [2]
 RP NUCLEOTIDE SEQUENCE (ISOFORM 1).
 RC TISSUE=Lung;
 RX MEDLINE=98445416; PubMed=9770531; DOI=10.1073/pnas.95.21.12596;
 RA Ivanov S.V., Kuzmin I., Wei M.-H., Pack S., Geil L., Johnson B.E.,
 RA Stanbridge E.J., Lerman M.I.;
 RT "Down-regulation of transmembrane carbonic anhydriases in renal cell
 RT carcinoma cell lines by wild-type von Hippel-Lindau transgenes.";
 RL Proc. Natl. Acad. Sci. U.S.A. 95:12596-12601(1998).
 RN [3]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORMS 1 AND 2).
 RC TISSUE=Eye, and Kidney;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altshul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny B., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [4]
 RP X-RAY CRYSTALLOGRAPHY (1.5 ANGSTROMS) OF 30-292.
 RX MEDLINE=21396545; PubMed=11493685; DOI=10.1073/pnas.161301298;
 RA Whittington D.A., Waheed A., Ulmasov B., Shah G.N., Grubb J.H.,
 RA Sly W.S., Christianson D.W.;
 RT "Crystal structure of the dimeric extracellular domain of human
 RT carbonic anhydrase XII, a bitopic membrane protein overexpressed in
 RT certain cancer tumor cells.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:9545-9550(2001).
 CC -!- FUNCTION: Reversible hydration of carbon dioxide.
 CC -!- CATALYTIC ACTIVITY: H(2)CO(3) = CO(2) + H(2)O.

CC -!- COFACTOR: Zinc (By similarity).
CC -!- ENZYME REGULATION: Inhibited by acetazolamide.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=O43570-1; Sequence=Displayed;
CC Name=2;
CC IsoId=O43570-2; Sequence=VSP_000772;
CC Note=No experimental confirmation available;
CC TISSUE SPECIFICITY: Highly expressed in colon, kidney, prostate,
CC intestine and activated lymphocytes. Expressed at much higher
CC levels in the renal cell cancers than in surrounding normal kidney
CC tissue. Moderately expressed in pancreas, ovary and testis.
CC -!- SIMILARITY: Belongs to the eukaryotic-type carbonic anhydrase
CC family.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC
CC ENBL; AF051882; AAC39789.1; -; mRNA.
CC ENBL; AF037335; AAC63952.1; -; mRNA.
CC ENBL; BC000278; AAH00278.1; -; mRNA.
CC ENBL; BC011691; AAH11691.1; -; mRNA.
CC ENBL; BC023981; AAH23981.1; -; mRNA.
CC PDB; 1JC2; X-ray; A/B=30-291.
CC PDB; 1JDO; X-ray; A/B=30-291.
CC Ensemble; ENSG00000074410; Homo sapiens.
CC HGNC; HGNC:1371; CA12.
CC H-InvDB; HIX0012325; -.
CC MIM; 603263; -.
CC GO; GO:0016021; C:integral to membrane; TAS.
CC GO; GO:0004089; F:carbonate dehydratase activity; TAS.
CC GO; GO:0008270; F:zinc ion binding; TAS.
CC InterPro; IPR001148; Euk Coanhd; 1.
CC PANTHER; PTHR18952; Euk Coanhd; 1.
CC Pfam; PF00194; Carb anhydrase; 1.
CC PRODOM; PD000865; Euk Coanhd; 1.
CC PROSITE; PS00162; EUK_CO2_ANHYDRASE; 1.
CC 3D-structure; Alternative splicing; Glycoprotein; Lyase;
CC Metal-binding; Signal; Transmembrane; Zinc.
CC SIGNAL 1 24 Potential.
CC CHAIN 25 354 Carbonic anhydrase XII.
CC TOPO_DOM 25 301 Extracellular (Potential).
CC TRANSMEM 302 322 Potential.
CC TOPO_DOM 323 354 Cytoplasmic (Potential).
CC METAL 119 119 Zinc (catalytic).
CC METAL 121 121 Zinc (catalytic).
CC METAL 145 145 Zinc (catalytic).
CC CARBOHYD 28 28 N-linked (GlcNAc...) (Potential).
CC CARBOHYD 80 80 N-linked (GlcNAc...) (Potential).
CC CARBOHYD 162 162 N-linked (GlcNAc...) (Potential).
CC DISULFID 50 230 Missing (in isoform 2).
CC VARSPLIC 292 302 /FTid=VSP_000772.
CC
CC TURN 37 38
CC HELIX 40 42
CC HELIX 43 46
CC HELIX 48 51
CC STRAND 59 60
CC HELIX 62 64
CC STRAND 65 67
CC TURN 69 70
CC STRAND 75 77
CC STRAND 80 80
CC TURN 83 84
CC STRAND 86 91
CC STRAND 96 99
CC TURN 102 103
CC STRAND 105 107

113 122
124 124
126 126
127 127
134 135
136 137
138 138
142 150
151 153
157 160
161 162
164 165
167 176
181 183
184 187
188 189
190 192
196 197
199 203
206 206
207 210
215 216
218 223
228 229
234 239
243 245
247 255
258 258
262 263
269 269
285 287
SQ SEQUENCE 354 AA; 39451 MW; 9016216BF2CA6C0C CRC64;

Query Match 22.2%; Score 537.5; DB 1; Length 354;
Best Local Similarity 35.8%; Pred. No. 9.2e-24;
Matches 129; Conservative 46; Mismatches 142; Indels 43; Gaps 10;

QY 117 EAPGDQEPONNAHRDKGDDQSHWRY---GGPPMPVRVSPACAGRFQSPVDIRPQLAAAF 173
Db 19 EQSSP-APVNG-----SKWTFGPDGNSMSKKYPSCGGLQSPIDLHSDILQY 67

QY 174 CPALRPLELGLFQPLPELRNRNGSHSVOLTTPGLEMALPGREYRALQLHLHWAGAG 233
Db 68 DASITPLEFQGYNLANKQFLLTNGHVKLNLPDMDHIQ-GLQSRYSATQLHLHWGNPN 126

QY 234 RP-GSEHTVEGHRPAPAIHVHVL-STAFARVDEALGRPGGLAVLAAFLLEGPEENSAYEQ 291
Db 127 DPHGSEHTVSGQHFAAEHLHVYNSDLYPDASTASNKSEGLAVLAVLEMG-SFNPSYDK 185

QY 292 LLSPLLEIAEGSETQVPGDIDISALLPSDFSRYPQYEGSLTTPPCAQGVITVFNQTVML 351
Db 186 IFSHLQHVKYKGQAFVPGFNIELLPERTAEYRYRGSLLTTPCNPVLVTVFRNPVQI 245

QY 352 SAKQLHTLSDTLW-----GFGDSRLQLNFRATQPLNGRVIEASPPAGVDSPPRAEPVOL 406
Db 246 SQEQLLALETALYCTHMDDFSPREMINNFRQVKQDERLVVTSFS-----QV 292

QY 407 NSCLAAGDIILAVFGLLFA-----VTSVAFVQMRQHRGTGKGVSYRPAEVAETGA 459
Db 293 QVCTAAGLSLGIILSLALAGILGICIVVSVVSIWLFRRKSIKKGDNKGVYKRPATOMETEA 352

RESULT 8
CAH6 SHEEP
ID CAH6 SHEEP STANDARD; PRT; 307 AA.
AC P08060;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Carbonic anhydrase VI (EC 4.2.1.1) (Carbonate dehydratase VI) (CA-VI)
DE (Secreted carbonic anhydrase) (Salivary carbonic anhydrase).
GN Name=CA6;
OS Ovis aries (Sheep).

OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Laurasiatida; Cetartiodactyla; Ruminantia;
OC	Pecora; Bovidae; Caprinae; Ovis.
NCBI_TaxID=9940;	
FN	[1]
RP	PROTEIN SEQUENCE.
RC	TISSUE=Saliva;
RX	MEDLINE=38294021; PubMed=3135834;
RX	Fernley R.T., Wright R.D., Coghlan J.P.;
RT	"Complete amino acid sequence of ovine salivary carbonic anhydrase.";
RL	Biochemistry 27:2815-2820(1988).
CC	-1- FUNCTION: Reversible hydration of carbon dioxide. Its role in
CC	saliva is unknown.
CC	-1- CATALYTIC ACTIVITY: H(2)CO(3) = CO(2) + H(2)O.
CC	-1- COFACTOR: Zinc (By similarity).
CC	-1- SUBCELLULAR LOCATION: Secreted.
CC	-1- TISSUE SPECIFICITY: Major constituent of saliva.
CC	-1- SIMILARITY: Belongs to the eukaryotic-type carbonic anhydrase
CC	family.
CC	-----
CC	This Swiss-Prot entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use as long as its content is in no way modified and this statement is not
CC	removed.
CC	-----
DR	PIR; A29393; A29393.
DR	HSP; O43570; IUD0.
DR	InterPro; IPR011148; Euk_Coanhd.
DR	PANTHER; PTHR18952; Euk_Coanhd; 1.
DR	Pfam; PF00194; Carb anhydrase; 1.
DR	ProDom; PD000865; Euk_Coanhd; 1.
DR	PROSITE; PS00162; Euk_CO2_ANGYDRASE; 1.
KW	Direct protein sequencing; Glycoprotein; Lyase; Metal-binding; Zinc.
FT	METAL 94 94
FT	METAL 96 96
FT	METAL 121 121
FT	CARBOHYD 50 50
FT	CARBOHYD 239 239
FT	DISULFID 25 207
FT	VARIANT 63 63
FT	VARIANT 297 297
FT	SEQUENCE 307 AA; 35555 MW; 338682C2D45E5D6C CRC64;
SQ	

Query Match	21.7%	Score 526;	DB 1;	Length 307;
Best Local Similarity	43.2%	Pred. No. 3.7e-23;		
Matches 111: Conservative	40:	Mismatches 98:	Indels	8:
			Gaps	4:

Qy	141	WRYG----	GDPPMPRVSPACAGAFQSPVDITRPQLAAFCPALRPLELLGLGFOLPPELURLR	196
Db	6	WTYSEGLMDEAHMPL	LEYPKCGGRQSPIDLQMKKKVQYNPSLRALNTGYGLWH-GEFPVPT	64
Qy	197	NNGHSVOLTLP	PGLEMALGPREYARLQLHLHWAAGR--PGSEHTVEGHRFPAETHVVH	254
Db	65	NNGHTVQISL	PSTMTTSDGTQYLAQMHPHGGASSEISGSEHTVDGNRYVIEIHVVH	124
Qy	255	LSTAFARVDALGR	PGGLAVLAAPLE-EGPEENSAYEQLLSRLLEETAEBSGETQVPGLDI	313
Db	125	YNSKYNSEYEAQKEP	PDGLAVLAALVEKDYENAYYSKFISHLEDIRYAGQSTVLRGLDI	184
Qy	314	SALLPSPFSRYFOYEGSLTTP	PPCAQGYIWWFNTQVMLSAQLHLTSDTLWPGDSDRLQL	373
Db	185	EDMLPGDLRYYSYLSGLTTP	PCCTENVHFWFVADTVKLSKTQVEKLENSLLNHQNTIQN	244
Qy	374	NFRATOP	NGRVIEASF	390
Db	245	DYRRTOP	LNHRVVEANF	261

RESULT 9
Q53YE5 HUMAN
ID Q53YE5 HUMAN PRELIMINARY; PRT; 343 AA.
AC Q53YE5;

	13-SEP-2005	(TrEMBLrel. 31, Created)	
DT	13-SEP-2005	(TrEMBLrel. 31, Last sequence update)	
DT	13-SEP-2005	(TrEMBLrel. 31, Last annotation update)	
DT	13-SEP-2005	(TrEMBLrel. 31, Last annotation update)	
OS	Carbonic anhydrase XII.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;		
OC	Homo.		
NCBI_TaxID=9606;			
[1]			
RP	NUCLEOTIDE SEQUENCE.		
RA	Kaline N., Chen X., Rolfs A., Halleck A., Hines L., Eisenstein S.,		
RA	Koundinya M., Raphael J., Moreira D., Kelley T., Labaer J., Lin Y.,		
RA	Prelan M., Farmer A.,		
RT	"Cloning of human full-length cDNAs in BD Creator(TM) System Donor		
RT	vector."		
RL	Submitted (MAY-2003) to the EMBL/GenBank/DDBJ databases.		
DR	EMBL; BT006656; AAP35302.1; -; mRNA.		
SQ	SEQUENCE 343 AA; 38407 MW; 87FD35255B37D18 CRC64;		
	Query Match	21.7%; Score 525; DB 2; Length 343;	
	Best Local Similarity	36.0%; Pred. No. 4.8e-23;	
	Matches 127; Conservative	45; Mismatches 141; Indels 40; Gaps 10;	
Qy	117	EAPGDPOEPQNNAHRDKEGDDQSHWRY--GSDPPMPVRSPACAGRFQSPVDIRPQLAAAF 173	
Dd	19	EQFSSP-APVNG-----SKWTYFGPDGENSMKSKYPSCGLGLQSLDLHSDLIQY 67	
Qy	174	CPALRPLELLGFQPLPELRNRNNGHSVOLTPPGLEMALGPGRYRALQLHLHWGAAG 233	
Dd	68	DASLTPLFEQGYNLNSANKQFLLTNNGHSVKLNI-PDSMHIO-GLOSRYSATQLHLHWGNPN 126	
Qy	234	RP-GSEHTVGHFPAEIHVVHL-S'AFARVDEALGRPGLAVLAAPLESGPENSAYEQ 291	
Dd	127	DPHGSEHTVSGQHFAELHIVHNSDLYPDASTASNKSEGVLAVLIENG-SFNPSYDK 185	
Qy	292	LLSRLBEIAEEGSETQVPGLDISALLPSDFSRFYQEGSLTTPPCAQGVITWTFNQTVML 351	
Dd	186	IFSHLOHVKYKGQEAFFVPGFNIELLPARTAEAYRYRGSLTTPPCNPVTIWLTVFRNPVQI 245	
Qy	352	SAKOLHTLSDTLM-----GPGDSRLQNRNPTOPLNGRVIEAASFPGADVSSPRAAEFPVL 406	
Dd	246	SQSQQLLALETALYCTHMDDPSREMIINNFQVOKFDERLVITYSFSQGIILS----- 296	
Qy	407	NSCLAAAGDIIALVFGLLFAYTSVAFLVQMRRQRHRRGTGGVSYRPAEVAFTGA 459	
Dd	297	---LALAGILGIC-----VVVYSIWLFRKSTIKKGNKGVIYKPAFKMETEA 341	

RESULT 10

Q6GQ30_XENLA Q6GQ30_XENLA PRELIMINARY; PRT; 525 AA.
ID Q6GQ30;
AC Q6GQ30;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE MG80389 protein.
DE Name=MG80389;
GN Name=Xenopus laevis (African clawed frog).
OS Xenopus laevis (African clawed frog).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
OC Xenopodinae; Xenopus;
RX NCBI_TaxID=8355;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Liver;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altshul S.F., Zengberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.G., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,

RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Ketterman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Marra M.A.,
 RA Schnurch A., Schein J.E., Jones S.J.M., Marra M.A.,
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 [2]
 RN NUCLEOTIDE SEQUENCE.
 RP TISSUE=Liver;
 RC MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
 RX Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
 RA "Genetic and genomic tools for Xenopus research: The NIH Xenopus
 RT initiative.";
 RT Dev. Dyn. 225:384-391(2002).
 [3]
 RN NUCLEOTIDE SEQUENCE.
 RP TISSUE=Liver;
 RC Klein S., Gerhard D.S.;
 RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
 CC -I- COPACTOR: Binds 2 calcium ions per subunit (By similarity).
 CC -I- SUBCELLULAR LOCATION: Secreted (By similarity).
 DR EMBL; BC072918; AAH72918.1; -, mRNA.
 DR GO; GO:0004089; F:carbonate dehydratase activity; IEA.
 DR GO; GO:0008270; F:zinc ion binding; IEA.
 DR GO; GO:0006730; P:one-carbon compound metabolism; IEA.
 DR InterPro; IPR001148; Euk Coanhd.
 DR InterPro; IPR001759; Pentaxin.
 DR Pfam; PF00194; Carb_anhydase; 1.
 DR Pfam; PF00354; Pentaxin; 1.
 DR PRINTS; PR00895; PENTAXIN.
 DR ProDom; PD000865; Euk Coanhd; 1.
 DR ProDom; PD002153; Pentaxin; 1.
 DR SMART; SM00359; PTX; 1.
 DR PROSITE; PS00162; EUK_CO2_ ANHYDRASE; 1.
 DR PROSITE; PS00289; PENTAXIN; UNKNOWN_1.
 KW Pentaxin.
 SQ SEQUENCE 525 AA; 59644 MW; DIDC237D63735578 CRC64;
 Query Match 21.4%; Score 519; DB 2; Length 525;
 Best Local Similarity 40.8%; Pred. No. 1.8e-22;
 Matches 107; Conservative 47; Mismatches 98; Indels 10; Gaps 4;
 QY 139 SH---WRYG---GDPWPVRVSPACAGRFQSPVDIRPQLAAFCPALRPLELLGFLPLP 191
 DB 22 SHVETVYQEGELDEANWKKYPTCAKHQSPIDIQRKVRHNPQLQLIGYDGLTG 81
 QY 192 ELRLRNNGHSVQLTLPFCLEMALGPGREYRALQHLHWA--AGRPGSEHTVEGHRPPAE 249
 DB 82 HFKVTNNGHSVQIDLPSTMTIKGLNSLYTAVQMLHWGGLSESTGSEHTIDGMYLAE 141
 QY 250 IHVHL--STAPARVDEALGRPGGLAVLAFLFEEGPEENSAVEQLLSLEETAESESTOV 308
 DB 142 LHVHNSGAYKSFDEAKDKENGLAVLAFLYTTNGYNTYTSDFISKIAKIRYAGQETEM 201
 QY 309 PGLDISALLPDSFRYFQYEGSLTTPPCAQGVITWVFNQTMVLSAKQLHTLSDTLWPGD 368
 DB 202 HTLDVMAMLPENLVNFRYDGLSTTPPCTENVLWTFDPSVFLSKTQIDLLLENTLLDWHN 261
 QY 369 SRLQLNFRATQPLNGRVEASF 390
 DB 262 KTLANDYRHAQPHDRTVEASF 283

RESULT 11

CAH6 CANFA STANDARD; PRT; 320 AA.
 ID Q865C0;
 AC 10-MAY-2005 (Rel. 47, Created)
 DT 10-MAY-2005 (Rel. 47, Last sequence update)
 DT 13-SEP-2005 (Rel. 48, Last annotation update)
 DE Carbonic anhydrase VI precursor (EC 4.2.1.1) (Carbonate dehydratase VI) (CA-VI).
 GN Name-CA6;
 OS Canis familiaris (Dog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Canidae; Canis.
 OC NCBI_TaxID=9615;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [MRNA].
 RC TISSUE=Parotid gland;
 RA Murakami M.;
 RT "Canine carbonic anhydrase VI (CA6), mRNA.";
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 CC -I- FUNCTION: Reversible hydration of carbon dioxide. Its role in saliva is unknown (By similarity).
 CC -I- CATALYTIC ACTIVITY: H(2)CO(3) = CO(2) + H(2)O.
 CC -I- COPACTOR: Zinc (By similarity).
 CC -I- SUBCELLULAR LOCATION: Secreted.
 CC -I- SIMILARITY: Belongs to the eukaryotic-type carbonic anhydrase family.

 CC This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use as long as its content is in no way modified and this statement is not removed.

 DR EMBL; AB080972; BAC65098.1; -, mRNA.
 DR HSSP; O43570; IJDO.
 DR Ensembl; ENSCAPG0000019708; Canis familiaris.
 DR InterPro; IPR001148; Euk Coanhd.
 DR PANTHER; PTHR18952; Euk Coanhd; 1.
 DR Pfam; PF00194; Carb_anhydase; 1.
 DR ProDom; PD000865; Euk Coanhd; 1.
 DR PROSITE; PS00162; EUK_CO2_ ANHYDRASE; 1.
 DR Glycoprotein; Lyase; Metal-binding; Signal; Zinc.
 FT SIGNAL 1 17 Potential.
 FT CHAIN 18 320 Carbonic anhydrase VI.
 FT METAL 111 111 Zinc (catalytic) (By similarity).
 FT METAL 113 113 Zinc (catalytic) (By similarity).
 FT METAL 138 138 Zinc (catalytic) (By similarity).
 FT CARBOHYD 256 256 N-linked (GlcNAc...) (Potential).
 FT DISULFID 42 224 Potential.
 SQ SEQUENCE 320 AA; 36705 MW; 6E00A8D08AA0D23F CRC64;
 Query Match 21.2%; Score 515; DB 1; Length 320;
 Best Local Similarity 41.6%; Pred. No. 1.7e-22;
 Matches 107; Conservative 43; Mismatches 95; Indels 12; Gaps 4;
 QY 137 DOSHWRYGDPWPVRVSPACAGRFQSPVDIRPQLAAFCPALRPLELLGFLPLPLRLR 196
 DB 31 DQVH-----WPREYTCGTRQSPIDIDRKRKQVNPSPKALKLTGYRI-QVGEFPMI 81
 QY 197 NNGHSVQLTLPFCLEMALGPGREYRALQHLHWAAGR--PGSEHTVEGHRFPFAEIHVVH 254
 DB 82 NNGHTVQISLPPTWRWMSDGTVEIAQMHFWGSGASSEISGSEHTIDGIRFAEIHIVH 141
 QY 255 LSTAFARVDEALGRPGGLAVLAFLFEEGPEENSAVEQLLSLEETAESESTQVPGLDI 313
 DB 142 YNSKYSVDIAQHPDGLAVLAFLVQVDEYGYNTYNSFNFIHLNNIRYPGOSTVLSGLDI 201
 QY 314 SALLPDSFRYFQYEGSLTTPPCAQGVITWVFNQTMVLSAKQLHTLSDTLWPGDSRLQL 373
 DB 202 EDMLPENTHYTYRGSITTPPCTENVHFWLVHVRVLUSSIQTKWLENSILDHQNKLHS 261
 QY 374 NFRATQPLNGRVEASF 390

DR	EMBL; X96503; CAA65357.1; -; mRNA.	
DR	PIR; S71877; S71877.	
DR	HSSP; O43570; 1GD0.	
DR	InterPro; IPR001148; Euk COanhd.	
DR	PANTHER; PTHR18952; Euk COanhd; 1.	
DR	Pfam; PF00194; Carb anhydrase; 1.	
DR	ProDom; PD000865; Euk COanhd; 1.	
DR	PROSITE; PS00162; EUK_CO2_ANHYDRASE; 1.	
DR	Direct protein sequencing; Glycoprotein; Lyase; Metal-binding; Signal;	
DR	zinc.	
DR	SIGNAL	1
DR	CHAIN	15
DR	FT	14
DR	FT	15
DR	FT	319
DR	METAL	106
DR	FT	106
DR	METAL	108
DR	FT	108
DR	METAL	133
DR	FT	133
DR	CARBOHYD	62
DR	FT	62
DR	CARBOHYD	251
DR	FT	251
DR	DISULFID	37
DR	FT	219
DR	CONFLICT	16
DR	FT	16
DR	SEQUENCE	319 AA; 37007 MW; 179884A7A9083AED CRC64;
DR	Query Match	21.1%; Score 511; DB 1; Length 319;
DR	Best Local Similarity	42.0%; Pred. No.3e-22;
DR	Matches 108; Conservative 40; Mismatches 97; Indels 12; Gaps 4;	
DR	QY	137 DQSHRYGDDPWPRVSPACAGRFQSPDIRPOLAAFCALRPLLELLGOLFPLPELRLR 196
DR	DB	137 DQSHRYGDDPWPRVSPACAGRFQSPDIRPOLAAFCALRPLLELLGOLFPLPELRLR 196
DR	QY	26 DEKHWRLE-----QYPCGGTGTROSPIDLKMKKRYNPSLRALNLTGYGLRQ-GEFPMT 76
DR	DB	26 DEKHWRLE-----QYPCGGTGTROSPIDLKMKKRYNPSLRALNLTGYGLRQ-GEFPMT 76
DR	QY	197 NNGHSVQLTLPPLGLEMALPGREYRALQLHLHWG--AAGRPGSEHTVEGHRFPAAIHVVH 254
DR	DB	197 NNGHSVQLTLPPLGLEMALPGREYRALQLHLHWG--AAGRPGSEHTVEGHRFPAAIHVVH 254
DR	QY	77 NNGHTVQISLPSMSMTTSDGSGQYLAKQMHFWGSDSSEISGSEHTVDGMRYIEIHVVH 136
DR	DB	77 NNGHTVQISLPSMSMTTSDGSGQYLAKQMHFWGSDSSEISGSEHTVDGMRYIEIHVVH 136
DR	QY	255 LSTAFARVDEALGRPGGLAVLAFLF-EGPEENSAYEQLLSRLLEETAESESTQVPGLDI 313
DR	DB	255 LSTAFARVDEALGRPGGLAVLAFLF-EGPEENSAYEQLLSRLLEETAESESTQVPGLDI 313
DR	QY	137 YHSKYGSYBEAQNEPDLGLAVLAALVEVKDYAENTYYSNFISHLEDIRYAGQSTVLRDLDI 196
DR	DB	137 YHSKYGSYBEAQNEPDLGLAVLAALVEVKDYAENTYYSNFISHLEDIRYAGQSTVLRDLDI 196
DR	QY	314 SALLPSDFSRYQYEGSLTTPCAQGVITVFNQTVMLSAKQLHTLSDTLWPGDSRLQL 373
DR	DB	314 SALLPSDFSRYQYEGSLTTPCAQGVITVFNQTVMLSAKQLHTLSDTLWPGDSRLQL 373
DR	QY	197 QDMLPGDLRYYSYLGSLTTPSCSTENVHVFVADTVKLSTQTEKLENSLLNHQNETION 256
DR	DB	197 QDMLPGDLRYYSYLGSLTTPSCSTENVHVFVADTVKLSTQTEKLENSLLNHQNETION 256
DR	QY	374 NFRATOPLNGRVIEASF 390
DR	DB	374 NFRATOPLNGRVIEASF 390
DR	QY	257 NYRSQTPLNHRVVEANF 273
DR	DB	257 NYRSQTPLNHRVVEANF 273
DR	RESULT 15	
DR	Q4RVRL_TETNG PRELIMINARY; PRT; 282 AA.	
DR	ID	Q4RVRL_TETNG PRELIMINARY; PRT; 282 AA.
DR	AC	Q4RVRL1;
DR	DT	13-SEP-2005 (TEMBLrel. 31, Created)
DR	DT	13-SEP-2005 (TEMBLrel. 31, Last sequence update)
DR	DT	13-SEP-2005 (TEMBLrel. 31, Last annotation update)
DR	DE	Chromosome 9 SCAF14991, whole genome shotgun sequence.
DR	DE	(Fragment)
DR	DE	ORFNames=GSTENG00028211001;
DR	GN	Tetradon nigroviridis (Green puffer).
DR	OS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
DR	OC	Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
DR	OC	Acanthomorpha; Acanthopterygii; Percormorpha; Tetraodontiformes;
DR	OC	Tetraodontidae; Tetraodontidae; Tetraodon.
DR	OC	NCBI_TaxID=99883;
DR	OC	[1]
DR	RN	NUCLEOTIDE SEQUENCE.
DR	RP	Jaillon O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,
DR	RA	Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
DR	RA	Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
DR	RA	Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
DR	RA	Anthouard V., Jubin C., Castelloli V., Katinka M., Vacherie B.,
DR	RA	Blemond C., Skalli Z., Coutanceau L., Poullain J., de Berardinis V.,
DR	RA	Craud C., Duprat S., Brottier P., Coutanceau J.P., Gouzy J.,
DR	RA	Parra G., Lardier G., Chappelle C., McKernan K.J., McEwan P., Bosak S.,
DR	RA	Kallis M., Wolff J.N., Guigo R., Zody M.C., Mesirov J.,

RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
RA Wincker P., Lander E.S., Weissbach J., Roest Crolius H.:
RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
RT the early vertebrate proto-karyotype.";
RL Nature 431:946-957(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RG Genoscope; Whitehead Institute Centre for Genome Research;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; CAAE01014991; CAG07521.1; -; Genomic_DNA.
FT NON TER 282 282
SQ SEQUENCE 282 AA; 32164 MW; B4E8C0176A094201 CRC64;

Query Match 20.9%; Score 507; DB 2; Length 282;
Best Local Similarity 39.8%; Pred. No. 4.4e-22;
Matches 104; Conservative 47; Mismatches 102; Indels 8; Gaps 4;

QY 137 DQSHWRYG---GDPWPVRVSPACAGRFQSPVDIRPQLAAFCPALRPLELLGFQLPPLPE 192
DB 22 DGIHWYKEGALDOMHWPTKYPCGKQKQSPIDIQORNVRFNPDMLQLSLSGYDAQOQT 81

QY 193 LRLRNGHVSQLTLPGLEWALQPGREYRALQLHLHGA--AGRPGSEHTVEGHRFPAB 250
DB 82 L-MTNNGHTVQIDLPPTVITEGLPGKYTAVQMHLHWGGWDLASGGKHTIDGVRVMAEL 140

QY 251 HVVHL-STAFARVDEALGRPGGLAVLAFLLEGPEENSAYEQLLSLEIEEGSETQVP 309
DB 141 HVVHNSDKYKSFIEARDKPDGLAVLAFFYDGHFENTYYSDFIANLGKIKYAGQSMYIS 200

QY 310 GLDISALLPDSFRYFYEGSLTPPCACQGIWTFVFNQTVMLSAKQLHTLSDTLWGP 369
DB 201 SLDVRSMLPENLHFRYQGSLLTPFCYESILWTFDPTLSHNIQIRKLESTLMDLDNK 260

QY 370 RLQLNFRATQPLNGRVIEASF 390
DB 261 TLWNDYRIAQPLNDRVVESSF 281

Search completed: February 12, 2006, 19:55:39
Job time : 249 secs

Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
1	2424	100.0	459	1	US-08-481-658B-2	Sequence 2, Appli	
2	2424	100.0	459	1	US-08-477-504A-2	Sequence 2, Appli	
3	2424	100.0	459	1	US-08-486-756A-2	Sequence 2, Appli	
4	2424	100.0	459	1	US-08-485-862B-2	Sequence 2, Appli	
5	2424	100.0	459	1	US-08-787-739-2	Sequence 2, Appli	
6	2424	100.0	459	2	US-08-487-077A-2	Sequence 2, Appli	
7	2424	100.0	459	2	US-08-485-863A-2	Sequence 2, Appli	
8	2424	100.0	459	2	US-08-485-049D-2	Sequence 2, Appli	
9	2424	100.0	459	2	US-09-178-115-2	Sequence 2, Appli	
10	2424	100.0	459	2	US-09-177-776-2	Sequence 2, Appli	
11	2424	100.0	459	2	US-09-772-719B-2	Sequence 2, Appli	
12	2424	100.0	459	2	US-08-260-190-6	Sequence 6, Appli	
13	2424	100.0	459	2	US-09-349-016-6042	Sequence 6042, Ap	
14	2420	99.8	473	2	US-09-349-016-10203	Sequence 2, Appli	
15	2227	91.9	422	2	US-08-335-469-2	Sequence 2, Appli	
16	2227	91.9	422	2	US-08-260-190-2	Sequence 87, Appli	
17	2007	82.8	377	2	US-08-787-739-87	Sequence 87, Appli	
18	2007	82.8	377	2	US-09-178-115-87	Sequence 87, Appli	
19	2007	82.8	377	2	US-09-177-776-87	Sequence 87, Appli	
20	1398.5	57.7	429	1	US-07-964-589-2	Sequence 2, Appli	
21	1398.5	57.7	429	4	PCT-US93-02024-2	Sequence 2, Appli	
22	1370	56.5	257	2	US-08-787-739-51	Sequence 51, Appli	
23	1370	56.5	257	2	US-09-178-115-51	Sequence 51, Appli	
24	1370	56.5	257	2	US-09-177-776-51	Sequence 51, Appli	
25	1364	56.3	256	1	US-08-481-658B-51	Sequence 51, Appli	
26	1364	56.3	256	1	US-08-477-504A-51	Sequence 51, Appli	
27	1364	56.3	256	1	US-08-486-756A-51	Sequence 51, Appli	

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/260,190
FILING DATE: 15-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Lauder, Leona L.
REGISTRATION NUMBER: 30,863
REFERENCE/DOCKET NUMBER: D-0021.3C
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-435-2034
TELEFAX: 415-435-0727
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 459 amino acids
TYPE: amino acid
STRANDEDNESS: linear
MOLECULE TYPE: protein
DESCRIPTION: First 37 amino acids represent
DESCRIPTION: signal peptide, and remaining amino acids
DESCRIPTION: represent mature protein
US-08-486-756A-2

Query Match 100.0%; Score 2424; DB 1; Length 459;
Best Local Similarity 100.0%; Pred. No. 1.3e-196;
Matches 459; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAPLCPSWLPPLIPAPAPGLTVQLLSLLLLMPVHPQRLPRMQEDSPGCGSGSGDDPL 60
DB 1 MAPLCPSWLPPLIPAPAPGLTVQLLSLLLLMPVHPQRLPRMQEDSPGCGSGSGDDPL 60
QY 61 GEEDLPSEEDSPREDDPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLP 120
DB 61 GEEDLPSEEDSPREDDPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLP 120
QY 121 DPQFPQNNNAHRDKGDDQSHWRYGGDPWPVRVSPACAGRFQSPVDIRPOLAFCPALRPL 180
DB 121 DPQFPQNNNAHRDKGDDQSHWRYGGDPWPVRVSPACAGRFQSPVDIRPOLAFCPALRPL 180
QY 181 ELLGFQPLPELRLRNNHSHVQLTLPGLMALPGREYRALQHLHLHWGAAGRPGSEHT 240
DB 181 ELLGFQPLPELRLRNNHSHVQLTLPGLMALPGREYRALQHLHLHWGAAGRPGSEHT 240
QY 241 VEGHRFPAAIHVVHLSTAFARVDEALGRPGGLAVLAFLLEGPEPNSAYEQLLSRLEIA 300
DB 241 VEGHRFPAAIHVVHLSTAFARVDEALGRPGGLAVLAFLLEGPEPNSAYEQLLSRLEIA 300
QY 301 EGGSTQVPGDLISALLPSDFSRYFQYEGSLTTPPCAQGVITWTFNQTVMLSAKQLHTLS 360
DB 301 EGGSTQVPGDLISALLPSDFSRYFQYEGSLTTPPCAQGVITWTFNQTVMLSAKQLHTLS 360
QY 361 DTLWGPDSRLQNLNFRATQPLNGRVIEASFAGVDSSPRAAEFVQVNSCLAAGDILALVF 420
DB 361 DTLWGPDSRLQNLNFRATQPLNGRVIEASFAGVDSSPRAAEFVQVNSCLAAGDILALVF 420
QY 421 GLLFAVTSVAFVQVNRHRRGRTKGGVSRPAEVAETGA 459
DB 421 GLLFAVTSVAFVQVNRHRRGRTKGGVSRPAEVAETGA 459

RESULT 4
US-08-485-862B-2
Sequence 2, Application US/08485862B
Patent No. 5989838
GENERAL INFORMATION:
APPLICANT: Zavada, Jan
APPLICANT: Pastorekova, Silvia
APPLICANT: Pastorek, Jaromir
TITLE OF INVENTION: MN Gene and Protein
NUMBER OF SEQUENCES: 86
CORRESPONDENCE ADDRESS:
ADDRESSEE: Leona L. Lauder
STREET: 6 Mariposa Court
CITY: Tiburon

TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-435-2034
TELEFAX: 415-435-0727
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 459 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
DESCRIPTION: First 37 amino acids represent
signal peptide, and remaining amino acids
represent mature protein
US-08-485-863A-2

Query Match 100.0%; Score 2424; DB 2; Length 459;
Best Local Similarity 100.0%; Pred. No. 1.3e-196;
Matches 459; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAPLCSPWLPPLIIPAPAGLTVQLLSLLLLMPVHPQRLPRMQEDSPGSGSGEDDPL 60
DB 1 MAPLCSPWLPPLIIPAPAGLTVQLLSLLLLMPVHPQRLPRMQEDSPGSGSGEDDPL 60
QY 61 GEEDLPSEEDSPREEDPPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLP 120
DB 61 GEEDLPSEEDSPREEDPPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLP 120
QY 121 DPQEPQNNARHDKGDDQSHWRYGGDPWPVRVSPACAGRFQSPVDIRPQLAAFCPALRPL 180
DB 121 DPQEPQNNARHDKGDDQSHWRYGGDPWPVRVSPACAGRFQSPVDIRPQLAAFCPALRPL 180
QY 181 ELLGFQPLPELRLNRNGHSVQLTLPGLLEMALPGREYRALQLHLHWAAGRPGSEHT 240
DB 181 ELLGFQPLPELRLNRNGHSVQLTLPGLLEMALPGREYRALQLHLHWAAGRPGSEHT 240
QY 241 VEGHRFPAAETHVHLSTAFARVDEALGRPGGLAVLAALFLEEGPEENSAAYEQLLSRLLEEIA 300
DB 241 VEGHRFPAAETHVHLSTAFARVDEALGRPGGLAVLAALFLEEGPEENSAAYEQLLSRLLEEIA 300
QY 301 EGGSETQVGLDISALLPSDFSRYFQYEGSLTTPPCAQGVITWVFNQTVMLSAKQLHTLS 360
DB 301 EGGSETQVGLDISALLPSDFSRYFQYEGSLTTPPCAQGVITWVFNQTVMLSAKQLHTLS 360
QY 361 DTLWPGDSRLQLNFRATQPLNGRVIEASFPAGVDSSPRAAEVQLNSCLAAGDILALVF 420
DB 361 DTLWPGDSRLQLNFRATQPLNGRVIEASFPAGVDSSPRAAEVQLNSCLAAGDILALVF 420
QY 421 GLLFAVTSVAFVQMRQRRGTGKGVSYRPAEVAETGA 459
DB 421 GLLFAVTSVAFVQMRQRRGTGKGVSYRPAEVAETGA 459

RESULT 8
US-08-485-049D-2
Sequence 2, Application US/08485049D
Patent No. 6204370
GENERAL INFORMATION:
APPLICANT: Zavada, Jan
APPLICANT: Pastorekova, Silvia
APPLICANT: Pastorek, Jaromir
TITLE OF INVENTION: MN Gene and Protein
NUMBER OF SEQUENCES: 86
CORRESPONDENCE ADDRESS:
ADDRESSEE: Leona L. Lauder
STREET: 369 Pine Street
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/485,863A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/260,190
FILING DATE: 15-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Lauder, Leona L.
REGISTRATION NUMBER: 30,863
REFERENCE/DOCKET NUMBER: D-0021.3G

; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION NUMBER: US/09/178,115
; APPLICATION NUMBER: US/08/485,049D
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/260,190
; FILING DATE: 15-JUN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lauder, Leona L.
; REGISTRATION NUMBER: 30,863
; REFERENCE/DOCKET NUMBER: D-0021.3E
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-981-2034
; TELEFAX: 415-981-0332
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 459 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; DESCRIPTION: First 37 amino acids represent
; DESCRIPTION: signal peptide, and remaining amino acids
; DESCRIPTION: represent mature protein
; US-08-485-049D-2

Query Match 100.0%; Score 2424; DB 2; Length 459;
Best Local Similarity 100.0%; Pred. No. 1.3e-196;
Matches 459; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MAPLCSPWLPPLIPAPAGLT	VQLLSLLLLMPVHPQRLPRMQEDSP	PLGGSSGGDDPL	60
DB	1	MAPLCSPWLPPLIPAPAGLT	VQLLSLLLLMPVHPQRLPRMQEDSP	PLGGSSGGDDPL	60
QY	61	GEEDLPSEEDSPREEDPPG	EEELPGBEDLPEVKPKSEEGSLKLED	LPTVEAPG	120
DB	61	GEEDLPSEEDSPREEDPPG	EEELPGBEDLPEVKPKSEEGSLKLED	LPTVEAPG	120
QY	121	DPOEPQNNAHKDEGDQSH	WRYGGDPWPVRVSPACAGRFQSPVD	IRPOLAAFCPALRPL	180
DB	121	DPOEPQNNAHKDEGDQSH	WRYGGDPWPVRVSPACAGRFQSPVD	IRPOLAAFCPALRPL	180
QY	181	ELLGFGOLPPLPELRLRN	NGHSVQLTLPGLLEMALGPGREYRALQLHLH	WGAAGRPGSEHT	240
DB	181	ELLGFGOLPPLPELRLRN	NGHSVQLTLPGLLEMALGPGREYRALQLHLH	WGAAGRPGSEHT	240
QY	241	VEGHRFPAEIHVVHLST	AFARVDEALGRPGGLAVLAAFL	EEGPEENSAYEQLLSRLEEIA	300
DB	241	VEGHRFPAEIHVVHLST	AFARVDEALGRPGGLAVLAAFL	EEGPEENSAYEQLLSRLEEIA	300
QY	301	EEGSETQVPGGLDISAL	LPSDFSRYFYEGSLTTPPCAQQGVITVFNQ	TWMLSAKQLHTLS	360
DB	301	EEGSETQVPGGLDISAL	LPSDFSRYFYEGSLTTPPCAQQGVITVFNQ	TWMLSAKQLHTLS	360
QY	361	DTLWGPDSRLQNLNFR	ATQPLNGRVIEASFPAGVDS	SPRAAEFVQVNSCLAAGDILALVF	420
DB	361	DTLWGPDSRLQNLNFR	ATQPLNGRVIEASFPAGVDS	SPRAAEFVQVNSCLAAGDILALVF	420
QY	421	GLLFAVTSVAFLVQMR	QRRRGTKGVSYPAEVAETGA	459	
DB	421	GLLFAVTSVAFLVQMR	QRRRGTKGVSYPAEVAETGA	459	

RESULT 9
US-09-178-115-2
; Sequence 2, Application US/09178115
; Patent No. 6297041
; GENERAL INFORMATION:
; APPLICANT: Zavada, Jan
; APPLICANT: Pastorekova, Silvia
; APPLICANT: Pastorek, Jaromir
; TITLE OF INVENTION: MN Gene and Protein

; FILE REFERENCE: D-0021.5A
; CURRENT APPLICATION NUMBER: US/09/178,115
; CURRENT FILING DATE: 1998-10-23
; EARLIER APPLICATION NUMBER: 09/177,776
; EARLIER FILING DATE: 1998-10-23
; EARLIER APPLICATION NUMBER: 08/787,739
; EARLIER FILING DATE: 1997-01-24
; EARLIER APPLICATION NUMBER: 08/485,049
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 08/486,756
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 08/477,504
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 08/481,658
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 08/485,862
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 08/485,863
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 08/487,077
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 08/260,190
; EARLIER FILING DATE: 1994-06-15
; EARLIER APPLICATION NUMBER: 08/177,093
; EARLIER FILING DATE: 1993-12-30
; EARLIER APPLICATION NUMBER: 07/964,589
; EARLIER FILING DATE: 1992-10-21
; EARLIER APPLICATION NUMBER: PV-709-92
; EARLIER FILING DATE: 1992-03-11
; NUMBER OF SEQ ID NOS: 116
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 459
; TYPE: PRT
; ORGANISM: HUMAN
; US-09-178-115-2

Query Match 100.0%; Score 2424; DB 2; Length 459;
Best Local Similarity 100.0%; Pred. No. 1.3e-196;
Matches 459; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MAPLCSPWLPPLIPAPAGLT	VQLLSLLLLMPVHPQRLPRMQEDSP	PLGGSSGGDDPL	60
DB	1	MAPLCSPWLPPLIPAPAGLT	VQLLSLLLLMPVHPQRLPRMQEDSP	PLGGSSGGDDPL	60
QY	61	GEEDLPSEEDSPREEDPPG	EEELPGBEDLPEVKPKSEEGSLKLED	LPTVEAPG	120
DB	61	GEEDLPSEEDSPREEDPPG	EEELPGBEDLPEVKPKSEEGSLKLED	LPTVEAPG	120
QY	121	DPOEPQNNAHKDEGDQSH	WRYGGDPWPVRVSPACAGRFQSPVD	IRPOLAAFCPALRPL	180
DB	121	DPOEPQNNAHKDEGDQSH	WRYGGDPWPVRVSPACAGRFQSPVD	IRPOLAAFCPALRPL	180
QY	181	ELLGFGOLPPLPELRLRN	NGHSVQLTLPGLLEMALGPGREYRALQLHLH	WGAAGRPGSEHT	240
DB	181	ELLGFGOLPPLPELRLRN	NGHSVQLTLPGLLEMALGPGREYRALQLHLH	WGAAGRPGSEHT	240
QY	241	VEGHRFPAEIHVVHLST	AFARVDEALGRPGGLAVLAAFL	EEGPEENSAYEQLLSRLEEIA	300
DB	241	VEGHRFPAEIHVVHLST	AFARVDEALGRPGGLAVLAAFL	EEGPEENSAYEQLLSRLEEIA	300
QY	301	EEGSETQVPGGLDISAL	LPSDFSRYFYEGSLTTPPCAQQGVITVFNQ	TWMLSAKQLHTLS	360
DB	301	EEGSETQVPGGLDISAL	LPSDFSRYFYEGSLTTPPCAQQGVITVFNQ	TWMLSAKQLHTLS	360
QY	361	DTLWGPDSRLQNLNFR	ATQPLNGRVIEASFPAGVDS	SPRAAEFVQVNSCLAAGDILALVF	420
DB	361	DTLWGPDSRLQNLNFR	ATQPLNGRVIEASFPAGVDS	SPRAAEFVQVNSCLAAGDILALVF	420
QY	421	GLLFAVTSVAFLVQMR	QRRRGTKGVSYPAEVAETGA	459	
DB	421	GLLFAVTSVAFLVQMR	QRRRGTKGVSYPAEVAETGA	459	

361 DTLWPGDSRLQINFRATQPLNGRVIEASFPAGVDSPPRAAEFVQLNSCIAAGDILALVF 420
361 DTLWPGDSRLQINFRATQPLNGRVIEASFPAGVDSPPRAAEFVQLNSCIAAGDILALVF 420
421 GLLFAVTSVAFVQMRQRRHRTKGGVSYPAEVAETGA 459
421 GLLFAVTSVAFVQMRQRRHRTKGGVSYPAEVAETGA 459

RESULT 11
US-09-772-719B-2
; Sequence 2, Application US/09772719B
; Patent No. 6770438
; GENERAL INFORMATION:
; APPLICANT: Zavada, Jan
; Pastorekova, Silvia
; Pastorek, Jaromir
; TITLE OF INVENTION: MN Gene and Protein
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leona L. Lauder
; STREET: 465 California Street, Suite 450
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30 (BPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09772,719B
; FILING DATE: 30-Jan-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/485,049
; FILING DATE: 07-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Lauder, Leona L.
; REGISTRATION NUMBER: 30,863
; REFERENCE/DOCKET NUMBER: D-0021.3A-2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-981-2034
; TELEFAX: 415-981-0332
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 459 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; DESCRIPTION: First 37 amino acids represent
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-772-719B-2

Query Match 100.0%; Score 2424; DB 2; Length 459;
Best Local Similarity 100.0%; Pred. No. 1.3e-196;
Matches 459; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAPLCPSFWLPLIPAPAGLTVQLLSLLLLMPVHPQRLPRMQEDSPGCGSGGDDPL 60
Db 1 MAPLCPSFWLPLIPAPAGLTVQLLSLLLLMPVHPQRLPRMQEDSPGCGSGGDDPL 60
Qy 61 GEEDLPSEEDSPREDDPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLP 120
Db 61 GEEDLPSEEDSPREDDPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLP 120
Qy 121 DPQSPQNNARHDKGDDQSHWRYGDDPPWPRVSPACAGRFQSPVDIRPOLAFCPALRPL 180
Db 121 DPQSPQNNARHDKGDDQSHWRYGDDPPWPRVSPACAGRFQSPVDIRPOLAFCPALRPL 180
Qy 181 ELLGFQPLPELRLNNGHVSQVLTLPGLMALPGREYRALQLHLHWGAAGRPGSEHT 240

Qy 361 DTLWPGDSRLQINFRATQPLNGRVIEASFPAGVDSPPRAAEFVQLNSCIAAGDILALVF 420
Db 361 DTLWPGDSRLQINFRATQPLNGRVIEASFPAGVDSPPRAAEFVQLNSCIAAGDILALVF 420
Qy 421 GLLFAVTSVAFVQMRQRRHRTKGGVSYPAEVAETGA 459
Db 421 GLLFAVTSVAFVQMRQRRHRTKGGVSYPAEVAETGA 459

RESULT 10
US-09-177-776-2
; Sequence 2, Application US/09177776A
; Patent No. 6297051
; GENERAL INFORMATION:
; APPLICANT: Zavada, Jan
; Pastorekova, Silvia
; Pastorek, Jaromir
; TITLE OF INVENTION: MN Gene and Protein
; FILE REFERENCE: D-0021.5A
; CURRENT APPLICATION NUMBER: US/09177,776A
; FILING DATE: 1998-10-23
; EARLIER APPLICATION NUMBER: 08/787,739
; EARLIER FILING DATE: 1997-01-24
; EARLIER APPLICATION NUMBER: 08/485,049
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 08/486,756
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 08/477,504
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 08/481,658
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 08/485,862
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 08/485,863
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 08/487,077
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 08/260,190
; EARLIER FILING DATE: 1994-06-15
; EARLIER APPLICATION NUMBER: 08/177,093
; EARLIER FILING DATE: 1993-12-30
; EARLIER APPLICATION NUMBER: 07/964,589
; EARLIER FILING DATE: 1992-10-21
; EARLIER APPLICATION NUMBER: PV-709-92
; EARLIER FILING DATE: 1992-03-11
; NUMBER OF SEQ ID NOS: 116
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 2
; LENGTH: 459
; TYPE: PRT
; ORGANISM: HUMAN
US-09-177-776-2

Query Match 100.0%; Score 2424; DB 2; Length 459;
Best Local Similarity 100.0%; Pred. No. 1.3e-196;
Matches 459; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAPLCPSFWLPLIPAPAGLTVQLLSLLLLMPVHPQRLPRMQEDSPGCGSGGDDPL 60
Db 1 MAPLCPSFWLPLIPAPAGLTVQLLSLLLLMPVHPQRLPRMQEDSPGCGSGGDDPL 60
Qy 61 GEEDLPSEEDSPREDDPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLP 120
Db 61 GEEDLPSEEDSPREDDPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLP 120
Qy 121 DPQSPQNNARHDKGDDQSHWRYGDDPPWPRVSPACAGRFQSPVDIRPOLAFCPALRPL 180
Db 121 DPQSPQNNARHDKGDDQSHWRYGDDPPWPRVSPACAGRFQSPVDIRPOLAFCPALRPL 180
Qy 181 ELLGFQPLPELRLNNGHVSQVLTLPGLMALPGREYRALQLHLHWGAAGRPGSEHT 240
Db 181 ELLGFQPLPELRLNNGHVSQVLTLPGLMALPGREYRALQLHLHWGAAGRPGSEHT 240
Qy 241 VEGHRFPAAEIHVHLSTAFARVDEALGRPGGLAVLAFLERGPENSAYEQLLSRLEBIA 300
Db 241 VEGHRFPAAEIHVHLSTAFARVDEALGRPGGLAVLAFLERGPENSAYEQLLSRLEBIA 300
Qy 301 EEGSETQVPGDLISALLSDFSRYFYQYEGSLTTPPCAGQVITWVFNQTVMLSAKQHTLS 360
Db 301 EEGSETQVPGDLISALLSDFSRYFYQYEGSLTTPPCAGQVITWVFNQTVMLSAKQHTLS 360

US-09-949-016-10203
; Sequence 10203, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR FILING DATE: 2000-04-14
; PRIOR FILING DATE: 2000-10-20
; PRIOR FILING DATE: 2000-10-20
; PRIOR FILING DATE: 2000-10-03
; PRIOR FILING DATE: 2000-10-03
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10203
; LENGTH: 473
; TYPE: PRT
; ORGANISM: Human
; US-09-949-016-10203

US-09-949-016-10203
; Sequence 10203, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR FILING DATE: 2000-04-14
; PRIOR FILING DATE: 2000-10-20
; PRIOR FILING DATE: 2000-10-20
; PRIOR FILING DATE: 2000-10-03
; PRIOR FILING DATE: 2000-10-03
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10203
; LENGTH: 473
; TYPE: PRT
; ORGANISM: Human
; US-09-949-016-10203

Query Match 99.8%; Score 2420; DB 2; Length 473;
Best Local Similarity 99.8%; Pred. No. 3e-196; Mismatches 0; Indels 0; Gaps 0;
Matches 458; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MAPLCPSPWLPLLPAPAGLTVQLLSLLLMFVHPQRLPMQEDSPPLGGSGGDDPL 60
Db 15 MAPLCPSPWLPLLPAPAGLTVQLLSLLLMFVHPQRLPMQEDSPPLGGSGGDDPL 74
Qy 61 GEEDLPSEEDSPREEDPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLP 120
Db 75 GEEDLPSEEDSPREEDPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLP 134
Qy 121 DPQFQNNNAHRDKEGDDQSHWRYGDDPPWPRVSPACAGRFQSPVDIRPOLAFCPALRPL 180
Db 135 DPQFQNNNAHRDKEGDDQSHWRYGDDPPWPRVSPACAGRFQSPVDIRPOLAFCPALRPL 194
Qy 181 ELLEGFQLPPLPELRLNNGHVSQVLTLPGLLEWALGPGREYRALQLHLHWGAAGRPGSEHT 240
Db 195 ELLEGFQLPPLPELRLNNGHVSQVLTLPGLLEWALGPGREYRALQLHLHWGAAGRPGSEHT 254
Qy 241 VEGHRFPAAEIHVHVLSTAFARVDEALGRPGGLAVLAALFLEGPENSAEYQQLLSRLEBIA 300
Db 255 VEGHRFPAAEIHVHVLSTAFARVDEALGRPGGLAVLAALFLEGPENSAEYQQLLSRLEBIA 314
Qy 301 EGGSETQVPGLDISALLPSDFSRYQYEGSLTTPPCAQGVITVFNQTVMLSAKQLHTLS 360
Db 315 EGGSETQVPGLDISALLPSDFSRYQYEGSLTTPPCAQGVITVFNQTVMLSAKQLHTLS 374
Qy 361 DTLWGPDSRLQNLNFRATQPLNGRVIEASFPAGVDSSPRAAEPVQLNSCLAAGDILALVF 420
Db 375 DTLWGPDSRLQNLNFRATQPLNGRVIEASFPAGVDSSPRAAEPVQLNSCLAAGDILALVF 434
Qy 421 GLLFAVTSVAFVQMRQRRGTGKGVSYRPAEVAETGA 459
Db 435 GLLFAVTSVAFVQMRQRRGTGKGVSYRPAEVAETGA 473

Query Match 99.8%; Score 2420; DB 2; Length 473;
Best Local Similarity 99.8%; Pred. No. 3e-196; Mismatches 0; Indels 0; Gaps 0;
Matches 458; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MAPLCPSPWLPLLPAPAGLTVQLLSLLLMFVHPQRLPMQEDSPPLGGSGGDDPL 60
Db 15 MAPLCPSPWLPLLPAPAGLTVQLLSLLLMFVHPQRLPMQEDSPPLGGSGGDDPL 74
Qy 61 GEEDLPSEEDSPREEDPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLP 120
Db 75 GEEDLPSEEDSPREEDPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLP 134
Qy 121 DPQFQNNNAHRDKEGDDQSHWRYGDDPPWPRVSPACAGRFQSPVDIRPOLAFCPALRPL 180
Db 135 DPQFQNNNAHRDKEGDDQSHWRYGDDPPWPRVSPACAGRFQSPVDIRPOLAFCPALRPL 194
Qy 181 ELLEGFQLPPLPELRLNNGHVSQVLTLPGLLEWALGPGREYRALQLHLHWGAAGRPGSEHT 240
Db 195 ELLEGFQLPPLPELRLNNGHVSQVLTLPGLLEWALGPGREYRALQLHLHWGAAGRPGSEHT 254
Qy 241 VEGHRFPAAEIHVHVLSTAFARVDEALGRPGGLAVLAALFLEGPENSAEYQQLLSRLEBIA 300
Db 255 VEGHRFPAAEIHVHVLSTAFARVDEALGRPGGLAVLAALFLEGPENSAEYQQLLSRLEBIA 314
Qy 301 EGGSETQVPGLDISALLPSDFSRYQYEGSLTTPPCAQGVITVFNQTVMLSAKQLHTLS 360
Db 315 EGGSETQVPGLDISALLPSDFSRYQYEGSLTTPPCAQGVITVFNQTVMLSAKQLHTLS 374
Qy 361 DTLWGPDSRLQNLNFRATQPLNGRVIEASFPAGVDSSPRAAEPVQLNSCLAAGDILALVF 420
Db 375 DTLWGPDSRLQNLNFRATQPLNGRVIEASFPAGVDSSPRAAEPVQLNSCLAAGDILALVF 434
Qy 421 GLLFAVTSVAFVQMRQRRGTGKGVSYRPAEVAETGA 459
Db 435 GLLFAVTSVAFVQMRQRRGTGKGVSYRPAEVAETGA 473

Query Match 91.9%; Score 2227; DB 2; Length 422;
Best Local Similarity 100.0%; Pred. No. 5.3e-180; Mismatches 0; Indels 0; Gaps 0;
Matches 422; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 38 ORLPRMOEDSPPLGGSGGDDPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLP 97
Db 1 ORLPRMOEDSPPLGGSGGDDPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLP 60
Qy 98 VKPKSEEGSLKLEDLPTVEAPGDPQFPQNNNAHRDKEGDDQSHWRYGDDPPWPRVSPACA 157
Db 61 VKPKSEEGSLKLEDLPTVEAPGDPQFPQNNNAHRDKEGDDQSHWRYGDDPPWPRVSPACA 120
Qy 158 GRFQSPVDIRPOLAFCPALRPLLELGFQPLPPLPELRLNNGHVSQVLTLPGLLEWALGPG 217
Db 121 GRFQSPVDIRPOLAFCPALRPLLELGFQPLPPLPELRLNNGHVSQVLTLPGLLEWALGPG 180
Qy 218 REYRALQLHLHWGAAGRPGSEHTVEGHRFPAAEIHVHVLSTAFARVDEALGRPGGLAVLAA 277
Db 181 REYRALQLHLHWGAAGRPGSEHTVEGHRFPAAEIHVHVLSTAFARVDEALGRPGGLAVLAA 240
Qy 278 FLEGPENSAEYQQLLSRLEBIAEGSETQVPGLDISALLPSDFSRYQYEGSLTTPPCA 337
Db 241 FLEGPENSAEYQQLLSRLEBIAEGSETQVPGLDISALLPSDFSRYQYEGSLTTPPCA 300
Qy 338 QGVITVFNQTVMLSAKQLHTLSDTLWGPDSRLQNLNFRATQPLNGRVIEASFPAGVDSS 397
Db 301 QGVITVFNQTVMLSAKQLHTLSDTLWGPDSRLQNLNFRATQPLNGRVIEASFPAGVDSS 360
Qy 398 PRAAEPVQLNSCLAAGDILALVFGLLFAVTSVAFVQMRQRRGTGKGVSYRPAEVAET 457
Db 361 PRAAEPVQLNSCLAAGDILALVFGLLFAVTSVAFVQMRQRRGTGKGVSYRPAEVAET 420
Qy 458 GA 459
Db 421 GA 422

Query Match 91.9%; Score 2227; DB 2; Length 422;
Best Local Similarity 100.0%; Pred. No. 5.3e-180; Mismatches 0; Indels 0; Gaps 0;
Matches 422; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 38 ORLPRMOEDSPPLGGSGGDDPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLP 97
Db 1 ORLPRMOEDSPPLGGSGGDDPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLP 60
Qy 98 VKPKSEEGSLKLEDLPTVEAPGDPQFPQNNNAHRDKEGDDQSHWRYGDDPPWPRVSPACA 157
Db 61 VKPKSEEGSLKLEDLPTVEAPGDPQFPQNNNAHRDKEGDDQSHWRYGDDPPWPRVSPACA 120
Qy 158 GRFQSPVDIRPOLAFCPALRPLLELGFQPLPPLPELRLNNGHVSQVLTLPGLLEWALGPG 217
Db 121 GRFQSPVDIRPOLAFCPALRPLLELGFQPLPPLPELRLNNGHVSQVLTLPGLLEWALGPG 180
Qy 218 REYRALQLHLHWGAAGRPGSEHTVEGHRFPAAEIHVHVLSTAFARVDEALGRPGGLAVLAA 277
Db 181 REYRALQLHLHWGAAGRPGSEHTVEGHRFPAAEIHVHVLSTAFARVDEALGRPGGLAVLAA 240
Qy 278 FLEGPENSAEYQQLLSRLEBIAEGSETQVPGLDISALLPSDFSRYQYEGSLTTPPCA 337
Db 241 FLEGPENSAEYQQLLSRLEBIAEGSETQVPGLDISALLPSDFSRYQYEGSLTTPPCA 300
Qy 338 QGVITVFNQTVMLSAKQLHTLSDTLWGPDSRLQNLNFRATQPLNGRVIEASFPAGVDSS 397
Db 301 QGVITVFNQTVMLSAKQLHTLSDTLWGPDSRLQNLNFRATQPLNGRVIEASFPAGVDSS 360
Qy 398 PRAAEPVQLNSCLAAGDILALVFGLLFAVTSVAFVQMRQRRGTGKGVSYRPAEVAET 457
Db 361 PRAAEPVQLNSCLAAGDILALVFGLLFAVTSVAFVQMRQRRGTGKGVSYRPAEVAET 420
Qy 458 GA 459
Db 421 GA 422

Search completed: February 12, 2006, 19:57:24
Job time : 50 secs

Search completed: February 12, 2006, 19:57:24
Job time : 50 secs

RESULT 15
US-08-335-469-2
; Sequence 2, Application US/08335469A
; Patent No. 6004535
; GENERAL INFORMATION:
; APPLICANT: Zavada, Jan
; APPLICANT: Pastorekova, Silvia
; APPLICANT: Pastorek, Jaromir
; TITLE OF INVENTION: Methods to Detect and Quantify MN Protein/Polypeptide
; TITLE OF INVENTION: Using MN-Specific Antibodies
; FILE REFERENCE: D-0021A
; CURRENT APPLICATION NUMBER: US/08/335,469A

RESULT 15
US-08-335-469-2
; Sequence 2, Application US/08335469A
; Patent No. 6004535
; GENERAL INFORMATION:
; APPLICANT: Zavada, Jan
; APPLICANT: Pastorekova, Silvia
; APPLICANT: Pastorek, Jaromir
; TITLE OF INVENTION: Methods to Detect and Quantify MN Protein/Polypeptide
; TITLE OF INVENTION: Using MN-Specific Antibodies
; FILE REFERENCE: D-0021A
; CURRENT APPLICATION NUMBER: US/08/335,469A

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 12, 2006, 20:07:51 ; Search time 177 Seconds

(without alignments)
1083.524 Million cell updates/sec

Title: US-09-967-237b-2

Perfect score: 2424

Sequence: 1 MAPLCSPWLLIPAPAG.....RRGKGGVSRPAEVAETGA 459

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA Main.*

- 1: /cgm2_6/ptodata/1/pubaa/US07_PUBCOMB.pep.*
- 2: /cgm2_6/ptodata/1/pubaa/US08_PUBCOMB.pep.*
- 3: /cgm2_6/ptodata/1/pubaa/US09_PUBCOMB.pep.*
- 4: /cgm2_6/ptodata/1/pubaa/US10A_PUBCOMB.pep.*
- 5: /cgm2_6/ptodata/1/pubaa/US10B_PUBCOMB.pep.*
- 6: /cgm2_6/ptodata/1/pubaa/US11_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2424	100.0	459	3	US-09-772-719-2
2	2424	100.0	459	3	Sequence 2, Appli
3	2424	100.0	459	4	Sequence 2, Appli
4	2424	100.0	459	4	Sequence 12, Appl
5	2424	100.0	459	4	Sequence 10, Appl
6	2424	100.0	459	4	Sequence 306, App
7	2424	100.0	459	4	Sequence 1239, Ap
8	2424	100.0	459	4	Sequence 16, Appli
9	2424	100.0	459	4	Sequence 2, Appli
10	2424	100.0	459	4	Sequence 138, App
11	2424	100.0	459	5	Sequence 461, App
12	2424	100.0	459	5	Sequence 6, Appli
13	2424	100.0	459	5	Sequence 2, Appli
14	2424	100.0	459	5	Sequence 19127, A
15	2424	100.0	459	5	Sequence 294, App
16	2424	100.0	459	5	Sequence 4846, Ap
17	2419	99.8	610	3	Sequence 1, Appli
18	2416	99.7	459	5	Sequence 92, Appl
19	2416	99.7	459	5	Sequence 92, Appl
20	2227	91.9	422	5	Sequence 2, Appli
21	2007	82.8	377	3	Sequence 87, Appl
22	2007	82.8	377	4	Sequence 7, Appli
23	2007	82.8	377	5	Sequence 5, Appli
24	1845	76.1	346	5	Sequence 104, App
25	1596	65.8	437	5	Sequence 90, Appl
26	1596	65.8	437	5	Sequence 18837, A
27	1587	65.5	437	5	Sequence 18838, A

Sequence 18868, A
Sequence 101, Appl
Sequence 51, Appl
Sequence 5, Appli
Sequence 9, Appli
Sequence 51, Appl
Sequence 92, Appl
Sequence 69, Appl
Sequence 105, Appl
Sequence 94, Appl
Sequence 54, Appl
Sequence 54, Appl
Sequence 423, App
Sequence 27, Appl
Sequence 74, Appl
Sequence 268, App
Sequence 74, Appl

ALIGNMENTS

RESULT 1

US-09-772-719-2
; Sequence 2, Application US/09772719
; Patent No. US20020137910A1
; GENERAL INFORMATION:
; APPLICANT: Zavada, Jan
; APPLICANT: Pastorekova, Silvia
; APPLICANT: Pastorek, Jaromir
; TITLE OF INVENTION: MN Gene and Protein
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leona L. Lauder
; STREET: 369 Pine Street
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA: US/09772,719
; APPLICATION NUMBER: US/09772,719
; FILING DATE: 30-JAN-2001
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/485,049
; FILING DATE: 07-JUN-1995
; APPLICATION NUMBER: US 08/260,190
; FILING DATE: 15-JUN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lauder, Leona L.
; REGISTRATION NUMBER: 30,863
; REFERENCE/DOCKET NUMBER: D-0021.3E
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-981-2034
; TELEFAX: 415-981-0332
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 459 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; DESCRIPTION: First 37 amino acids represent
; DESCRIPTION: signal peptide, and remaining amino acids
; DESCRIPTION: represent mature protein
US-09-772-719-2

Query Match		100.0%;	Score 2424;	DB 3;	Length 459;	
Best Local Similarity		100.0%;	Pred. No. 1.2e-173;			
Matches 459;		Conservative	0;	Mismatches	0;	Indels
QY	1	MAPLCPSPWPLLIIPAPAGLT	VQLLLSLLLLMPVHPQRLPRMQEDSP	PLGGSSGGDDPL	60	
DB	1	MAPLCPSPWPLLIIPAPAGLT	VQLLLSLLLLMPVHPQRLPRMQEDSP	PLGGSSGGDDPL	60	
QY	61	GBEDLPSEEDSPREEDPPG	EEEDLPGEEDLPGEEDLPGEEDLP	GEEDLPGEEDLP	PTVEAPG	120
DB	61	GBEDLPSEEDSPREEDPPG	EEEDLPGEEDLPGEEDLPGEEDLP	GEEDLPGEEDLP	PTVEAPG	120
QY	121	DPQEPQNNNAHRDKEGDDQ	SHWRYGGDPWPVRVSPACAGRFQSPVD	IRPQLAAFCPALRPL	180	
DB	121	DPQEPQNNNAHRDKEGDDQ	SHWRYGGDPWPVRVSPACAGRFQSPVD	IRPQLAAFCPALRPL	180	
QY	181	ELLGFGQLPPLPELR	LNNNGHSVQLTTPPGLEMALGPGREYRALQLHLH	WGAAGRPGSEHT	240	
DB	181	ELLGFGQLPPLPELR	LNNNGHSVQLTTPPGLEMALGPGREYRALQLHLH	WGAAGRPGSEHT	240	
QY	241	VEGHRFPAAETHVHVLSTAF	ARVDEALGRPGGLAVLAFLAFL	EEGPEENSAYEQLLSRLEEIA	300	
DB	241	VEGHRFPAAETHVHVLSTAF	ARVDEALGRPGGLAVLAFLAFL	EEGPEENSAYEQLLSRLEEIA	300	
QY	301	EGSETQVPGCLDISALLPS	DFSRYFOYEGSLTTPPCAQGVITVFNQT	VMLSAKQLHTLS	360	
DB	301	EGSETQVPGCLDISALLPS	DFSRYFOYEGSLTTPPCAQGVITVFNQT	VMLSAKQLHTLS	360	
QY	361	DTLWPGDLSRLQINFRATQ	PLNGRVIEASFPAGVDSSPRAAEPVQ	NSCLAAGDIILALVF	420	
DB	361	DTLWPGDLSRLQINFRATQ	PLNGRVIEASFPAGVDSSPRAAEPVQ	NSCLAAGDIILALVF	420	
QY	421	GLLFVAVTSVAFVQMRQ	RHRRGTGGVSYRPAEVAETGA	459		
DB	421	GLLFVAVTSVAFVQMRQ	RHRRGTGGVSYRPAEVAETGA	459		
RESULT 3						
US-10-301-822-12						
; Sequence 12, Application US/10301822						
; Publication No. US20030148410A1						
; GENERAL INFORMATION:						
; APPLICANT: Millennium Pharmaceuticals, Inc.						
; APPLICANT: Berger, Allison						
; APPLICANT: Guillemette, Tracy L.						
; APPLICANT: Kamatkar, Shubhangi						
; APPLICANT: Schlegel, Robert						
; APPLICANT: Monahan, John E.						
; APPLICANT: Thibodeau, Stephen N.						
; APPLICANT: BURGART, Lawrence J.						
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND						
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND						
; TITLE OF INVENTION: THERAPY OF COLON CANCER						
; FILE REFERENCE: MPM01-029P2RNM						
; CURRENT APPLICATION NUMBER: US/10/301,822						
; CURRENT FILING DATE: 2002-11-21						
; PRIOR APPLICATION NUMBER: US 60/339,971						
; PRIOR FILING DATE: 2001-12-10						
; PRIOR APPLICATION NUMBER: US 60/361,978						
; PRIOR FILING DATE: 2002-03-05						
; PRIOR APPLICATION NUMBER: US 60/381,988						
; PRIOR FILING DATE: 2002-05-20						
; NUMBER OF SEQ ID NOS: 228						
; SOFTWARE: FASTSEQ for Windows Version 4.0						
; SEQ ID NO 12						
; LENGTH: 459						
; TYPE: PRT						
; ORGANISM: Homo Sapiens						
US-10-301-822-12						
Query Match		100.0%;	Score 2424;	DB 4;	Length 459;	
Best Local Similarity		100.0%;	Pred. No. 1.2e-173;			
Matches 459;		Conservative	0;	Mismatches	0;	Indels
QY	1	MAPLCPSPWPLLIIPAPAGLT	VQLLLSLLLLMPVHPQRLPRMQEDSP	PLGGSSGGDDPL	60	
DB	1	MAPLCPSPWPLLIIPAPAGLT	VQLLLSLLLLMPVHPQRLPRMQEDSP	PLGGSSGGDDPL	60	
QY	61	GBEDLPSEEDSPREEDPPG	EEEDLPGEEDLPGEEDLPGEEDLP	GEEDLPGEEDLP	PTVEAPG	120
DB	61	GBEDLPSEEDSPREEDPPG	EEEDLPGEEDLPGEEDLPGEEDLP	GEEDLPGEEDLP	PTVEAPG	120
QY	121	DPQEPQNNNAHRDKEGDDQ	SHWRYGGDPWPVRVSPACAGRFQSPVD	IRPQLAAFCPALRPL	180	
DB	121	DPQEPQNNNAHRDKEGDDQ	SHWRYGGDPWPVRVSPACAGRFQSPVD	IRPQLAAFCPALRPL	180	
QY	181	ELLGFGQLPPLPELR	LNNNGHSVQLTTPPGLEMALGPGREYRALQLHLH	WGAAGRPGSEHT	240	
US-09-967-237-2						
; Sequence 2, Application US/09967237						
; Publication No. US20030049828A1						
; GENERAL INFORMATION:						
; APPLICANT: Zavada, Jan						
; APPLICANT: Pastorekova, Silvia						
; APPLICANT: Pastorek, Jaromir						
; TITLE OF INVENTION: MN Gene and Protein						
; FILE REFERENCE: D-0021-SB-2						
; CURRENT APPLICATION NUMBER: US/09/967,237						
; CURRENT FILING DATE: 2001-09-27						
; PRIOR APPLICATION NUMBER: 09/178,115						
; PRIOR FILING DATE: 1998-10-23						
; NUMBER OF SEQ ID NOS: 116						
; SOFTWARE: PatentIn Ver. 2.1						
; SEQ ID NO 2						
; LENGTH: 459						
; TYPE: PRT						
; ORGANISM: HUMAN						
US-09-967-237-2						
Query Match		100.0%;	Score 2424;	DB 3;	Length 459;	
Best Local Similarity		100.0%;	Pred. No. 1.2e-173;			
Matches 459;		Conservative	0;	Mismatches	0;	Indels
QY	1	MAPLCPSPWPLLIIPAPAGLT	VQLLLSLLLLMPVHPQRLPRMQEDSP	PLGGSSGGDDPL	60	
DB	1	MAPLCPSPWPLLIIPAPAGLT	VQLLLSLLLLMPVHPQRLPRMQEDSP	PLGGSSGGDDPL	60	
QY	61	GBEDLPSEEDSPREEDPPG	EEEDLPGEEDLPGEEDLPGEEDLP	GEEDLPGEEDLP	PTVEAPG	120
DB	61	GBEDLPSEEDSPREEDPPG	EEEDLPGEEDLPGEEDLPGEEDLP	GEEDLPGEEDLP	PTVEAPG	120
QY	121	DPQEPQNNNAHRDKEGDDQ	SHWRYGGDPWPVRVSPACAGRFQSPVD	IRPQLAAFCPALRPL	180	


```

; ORGANISM: HUMAN
US-10-795-933-6

Query Match      100.0%; Score 2424; DB 5; Length 459;
Best Local Similarity 100.0%; Pred. No. 1.2e-173;
Matches 459; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAPLCPSWMLPLLIIPAPAPGLTVQLLSLLLLMPVHPQRLPRMQEDSPPLGGSGSGEDDPL 60
   |||||
DB 1 MAPLCPSWMLPLLIIPAPAPGLTVQLLSLLLLMPVHPQRLPRMQEDSPPLGGSGSGEDDPL 60

QY 61 GEEDLPSEDSPREDDPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLP 120
   |||||
DB 61 GEEDLPSEDSPREDDPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLP 120

QY 121 DQEPQNNAHROKEDGQDQSHWRYGGDPPWPRVSPACAGRFQSPVDIRPOLAFCPALRPL 180
   |||||
DB 121 DQEPQNNAHROKEDGQDQSHWRYGGDPPWPRVSPACAGRFQSPVDIRPOLAFCPALRPL 180

QY 181 ELIGFQLPPLPELRLNNGHVSQVLTLPGLLEMALPGREYRALQLHLHWGAAGRPGESEHT 240
   |||||
DB 181 ELIGFQLPPLPELRLNNGHVSQVLTLPGLLEMALPGREYRALQLHLHWGAAGRPGESEHT 240

QY 241 VEGHRRPFAIHVVHLSTAFARVDEALGRPGGLAVLAFLLEGPEENSAYEQLLSRLBEIA 300
   |||||
DB 241 VEGHRRPFAIHVVHLSTAFARVDEALGRPGGLAVLAFLLEGPEENSAYEQLLSRLBEIA 300

QY 301 EGSETQVPGDLISALLPSDFSRFYQYEGSLTTPPCAQGVITWTFNQTVMLSAKQLHTLS 360
   |||||
DB 301 EGSETQVPGDLISALLPSDFSRFYQYEGSLTTPPCAQGVITWTFNQTVMLSAKQLHTLS 360

QY 361 DTLWPGDSRLQNLFRATQPLNGRVIEASFPAGVDSSPRAAEPVQLNSCLAAGDILALVF 420
   |||||
DB 361 DTLWPGDSRLQNLFRATQPLNGRVIEASFPAGVDSSPRAAEPVQLNSCLAAGDILALVF 420

QY 421 GLLFAVTSVAFLVQMRQRRRTGKGVSYRPAEVAETGA 459
   |||||
DB 421 GLLFAVTSVAFLVQMRQRRRTGKGVSYRPAEVAETGA 459

RESULT 12
US-10-888-694-2
; Sequence 2, Application US/10888694
; Publication No. US20050003425A1
; GENERAL INFORMATION:
; APPLICANT: Zavada, Jan
; Pastorekova, Silvia
; Pastorek, Jaromir
; TITLE OF INVENTION: MN Gene and Protein
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leona L. Lauder
; STREET: 465 California Street, Suite 450
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/888,694
; FILING DATE: 08-Jul-2004
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/772,719
; FILING DATE: 30-Jan-2001
; APPLICATION NUMBER: US 08/485,049
; FILING DATE: 07-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Lauder, Leona L.

```


REGISTRATION NUMBER: 30,863
REFERENCE/DOCKET NUMBER: D-0021.3A-2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-981-2034
TELEFAX: 415-981-0332
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 459 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: protein
DESCRIPTION: First 37 amino acids represent
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-888-694-2

Query Match 100.0%; Score 2424; DB 5; Length 459;
Best Local Similarity 100.0%; Pred. No. 1.2e-173;
Matches 459; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAPLCSPWMLPLIIPAPAGLTQVLLLSLLLLMPVHPQRLPRMQEDSPGSGSGDDPL 60
DB 1 MAPLCSPWMLPLIIPAPAGLTQVLLLSLLLLMPVHPQRLPRMQEDSPGSGSGDDPL 60
QY 61 GEEDLPSEEDSPREEDPPGEEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLP 120
DB 61 GEEDLPSEEDSPREEDPPGEEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLP 120
QY 121 DPOEPQNNARHDEKGGDQSHWRVGGDPMPVSPACAGRFQSPVDIRPQLAAFCPALRPL 180
DB 121 DPOEPQNNARHDEKGGDQSHWRVGGDPMPVSPACAGRFQSPVDIRPQLAAFCPALRPL 180
QY 181 ELGFGQLPPLPELRLRNNGHSVQLTLPGLLEMALPGREYRALQLHLHWAAGRPGSEHT 240
DB 181 ELGFGQLPPLPELRLRNNGHSVQLTLPGLLEMALPGREYRALQLHLHWAAGRPGSEHT 240
QY 241 VEGHRPPAEIHVVHLSTAFARVDEALGRPGGLAVLAFLAFLAFLAFLAFLAFLAFLA 300
DB 241 VEGHRPPAEIHVVHLSTAFARVDEALGRPGGLAVLAFLAFLAFLAFLAFLAFLAFLA 300
QY 301 EGSETQVPGDLISALLPSDFSYFYQEGSLTTPPCAQGVITVFNQVMSAKQLHTLS 360
DB 301 EGSETQVPGDLISALLPSDFSYFYQEGSLTTPPCAQGVITVFNQVMSAKQLHTLS 360
QY 361 DTLWPGDSRLQNFNFRATQPLNGRVIEASFPAGVDSPPRAAEPVQVNSCLAAGDILALVF 420
DB 361 DTLWPGDSRLQNFNFRATQPLNGRVIEASFPAGVDSPPRAAEPVQVNSCLAAGDILALVF 420
QY 421 GLLFAVTSVAFLVQMRQRHRRGTGGVSYRPAEVAETGA 459
DB 421 GLLFAVTSVAFLVQMRQRHRRGTGGVSYRPAEVAETGA 459

RESULT 13
US-10-921-590-2
Sequence 2, Application US/10921590
Publication No. US20050031623A1
GENERAL INFORMATION:
APPLICANT: Pastorek, Jaromir
APPLICANT: Zavada, Jan
APPLICANT: Ortova Gut, Marta
APPLICANT: Zatovicova, Miriam
APPLICANT: Pastorekova, Silvia
APPLICANT: Zavadova, Zuzanna
TITLE OF INVENTION: SOLUBLE FORM OF CARBONIC ANHYDRASE IX (s-CA IX), ASSAYS TO DETECT
TITLE OF INVENTION: s-CA IX, CA IX'S COEXPRESSION WITH HER-2/neu/c-erbB-2 AND CA
FILE REFERENCE: MX-2363/2376 US
CURRENT APPLICATION NUMBER: US/10/921,590
PRIOR FILING DATE: 2004-08-19
PRIOR FILING DATE: 2002-02-21
PRIOR APPLICATION NUMBER: 60/358,824
PRIOR APPLICATION NUMBER: 60/383,068

PRIOR FILING DATE: 2002-05-23
PRIOR APPLICATION NUMBER: 60/431,499
PRIOR FILING DATE: 2002-12-05
PRIOR APPLICATION NUMBER: PCT/US03/05136
PRIOR FILING DATE: 2003-02-21
PRIOR APPLICATION NUMBER: PCT/US03/05137
PRIOR FILING DATE: 2003-02-21
NUMBER OF SEQ ID NOS: 118
SOFTWARE: PatentIn version 3.2
SEQ ID NO 2
LENGTH: 459
TYPE: PRT
ORGANISM: Homo sapiens
US-10-921-590-2

Query Match 100.0%; Score 2424; DB 5; Length 459;
Best Local Similarity 100.0%; Pred. No. 1.2e-173;
Matches 459; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAPLCSPWMLPLIIPAPAGLTQVLLLSLLLLMPVHPQRLPRMQEDSPGSGSGDDPL 60
DB 1 MAPLCSPWMLPLIIPAPAGLTQVLLLSLLLLMPVHPQRLPRMQEDSPGSGSGDDPL 60
QY 61 GEEDLPSEEDSPREEDPPGEEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLP 120
DB 61 GEEDLPSEEDSPREEDPPGEEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLP 120
QY 121 DPOEPQNNARHDEKGGDQSHWRVGGDPMPVSPACAGRFQSPVDIRPQLAAFCPALRPL 180
DB 121 DPOEPQNNARHDEKGGDQSHWRVGGDPMPVSPACAGRFQSPVDIRPQLAAFCPALRPL 180
QY 181 ELGFGQLPPLPELRLRNNGHSVQLTLPGLLEMALPGREYRALQLHLHWAAGRPGSEHT 240
DB 181 ELGFGQLPPLPELRLRNNGHSVQLTLPGLLEMALPGREYRALQLHLHWAAGRPGSEHT 240
QY 241 VEGHRPPAEIHVVHLSTAFARVDEALGRPGGLAVLAFLAFLAFLAFLAFLAFLAFLA 300
DB 241 VEGHRPPAEIHVVHLSTAFARVDEALGRPGGLAVLAFLAFLAFLAFLAFLAFLAFLA 300
QY 301 EGSETQVPGDLISALLPSDFSYFYQEGSLTTPPCAQGVITVFNQVMSAKQLHTLS 360
DB 301 EGSETQVPGDLISALLPSDFSYFYQEGSLTTPPCAQGVITVFNQVMSAKQLHTLS 360
QY 361 DTLWPGDSRLQNFNFRATQPLNGRVIEASFPAGVDSPPRAAEPVQVNSCLAAGDILALVF 420
DB 361 DTLWPGDSRLQNFNFRATQPLNGRVIEASFPAGVDSPPRAAEPVQVNSCLAAGDILALVF 420
QY 421 GLLFAVTSVAFLVQMRQRHRRGTGGVSYRPAEVAETGA 459
DB 421 GLLFAVTSVAFLVQMRQRHRRGTGGVSYRPAEVAETGA 459

RESULT 14
US-10-732-923-19127
Sequence 19127, Application US/10732923
Publication No. US20050108791A1
GENERAL INFORMATION:
APPLICANT: Edgerton, Michael D
TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
FILE REFERENCE: 38-15(52796)C
CURRENT APPLICATION NUMBER: US/10/732,923
CURRENT FILING DATE: 2003-12-10
PRIOR APPLICATION NUMBER: 10/310,154
PRIOR FILING DATE: 2002-12-04
NUMBER OF SEQ ID NOS: 24149
SEQ ID NO 19127
LENGTH: 459
TYPE: PRT
ORGANISM: Homo sapiens
US-10-732-923-19127

Query Match 100.0%; Score 2424; DB 5; Length 459;
Best Local Similarity 100.0%; Pred. No. 1.2e-173;
Matches 459; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

	Matches	459;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
Qy	1	MAPLCPSFWLP	LPLLIPAPAGLTGVLQLLSLLMLPVHPQR	LRPMQEDSP	PLGGSSGEDDPL	60				
Db	1	MAPLCPSFWLP	LPLLIPAPAGLTGVLQLLSLLMLPVHPQR	LRPMQEDSP	PLGGSSGEDDPL	60				
Qy	61	GEEDLPSSEDSPREEDP	PGEEDLPGEEDLPVEVKPKSBEESGLKLEDLPTVPAEG	120						
Db	61	GEEDLPSSEDSPREEDP	PGEEDLPGEEDLPVEVKPKSBEESGLKLEDLPTVPAEG	120						
Qy	121	DPOBPQNNAURDKSGDDQSHWRYCGDPFPWPVRVS	PACAGRFOSPVDIRPOLAACFPCALRPL	180						
Db	121	DPOBPQNNAURDKSGDDQSHWRYCGDPFPWPVRVS	PACAGRFOSPVDIRPOLAACFPCALRPL	180						
Qy	181	ELIGCFOLPPPELRLRNNGHSVQLTPGLEMALPGGREYALOLHLHWGAAGRGSGEHT	240							
Db	181	ELIGCFOLPPPELRLRNNGHSVQLTPGLEMALPGGREYALOLHLHWGAAGRGSGEHT	240							
Qy	241	VEGRHFFAEIHVVHLSTAFARVDEALGRPGLAVLAFLAEEGPENSAYEQLLSRLSEIEIA	300							
Db	241	VEGRHFFAEIHVVHLSTAFARVDEALGRPGLAVLAFLAEEGPENSAYEQLLSRLSEIEIA	300							
Qy	301	BEGSETQVGLDISALLPSPDFSRFYQEGSLTTPPCAQGVITWFNQTVMLSAKQLHTLS	360							
Db	301	BEGSETQVGLDISALLPSPDFSRFYQEGSLTTPPCAQGVITWFNQTVMLSAKQLHTLS	360							
Qy	361	DTLWGPDGSRLQLNFRATQPLNGRVIEASPPAGVSDSPRAAPEVOLANSCLAGDILALVF	420							
Db	361	DTLWGPDGSRLQLNFRATQPLNGRVIEASPPAGVSDSPRAAPEVOLANSCLAGDILALVF	420							
Qy	421	GLLPFAVTSVAPLVQMRRQHRRGTCKGVSYPAAEVETGA	459							
Db	421	GLLPFAVTSVAPLVQMRRQHRRGTCKGVSYPAAEVETGA	459							

Search completed: February 12, 2006, 20:11:17
Job time : 179 secs

```

RESULT 15
US-10-794-514A-294
; Sequence 294, Application US/10794514A
; Publication No. US20050112134A1
; GENERAL INFORMATION:
; APPLICANT: Graddis, Thomas
; APPLICANT: Laus, Reiner
; APPLICANT: Diegel, Michael
; APPLICANT: Vidovic, Damir
; TITLE OF INVENTION: Compositions and Methods Employing Alternative
; TITLE OF INVENTION: Reading Frame Polypeptides for the Treatment of
; TITLE OF INVENTION: Cancer and Infectious Disease
; FILE REFERENCE: 11311.1003U
; CURRENT APPLICATION NUMBER: US/10/794,514A
; CURRENT FILING DATE: 2004-03-05
; NUMBER OF SEQ ID NOS: 733
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 294
; LENGTH: 459
; TYPE: PRT
; ORGANISM: Derived From Human Sequence
US-10-794-514A-294

Query Match          100.0%; Score 2424; DB 5; Length 459;
Best Local Similarity 100.0%; Pred. No. 1.2e-173;
Matches 459; Conservative 0; Mismatches 0; Indels 0; Gaps 0

Qy      1  MAPLCPSPWLPLLIIPAPAPGLTVQLLSLLLLMPVHPQRLPRMQEDSPPLGGSSGGDDPL 60
      1  MAPLCPSPWLPLLIIPAPAPGLTVQLLSLLLLMPVHPQRLPRMQEDSPPLGGSSGGDDPL 60

Qy      61  GREDLPSEEDSPREDDPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLP 120
      61  GREDLPSEEDSPREDDPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLP 120

Qy      121  DPQEPQNNAHNRKQGDGDQSHWRYGDDPPWPRVSPACAGRFQSPVDIRPQLAAFCPALRPL 180
      121  DPQEPQNNAHNRKQGDGDQSHWRYGDDPPWPRVSPACAGRFQSPVDIRPQLAAFCPALRPL 180

```

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 12, 2006, 20:08:26 ; Search time 17 Seconds
(without alignments)
354.309 Million cell updates/sec

Title: US-09-967-237B-2

Perfect score: 2424

Sequence: 1 MAPLCPSPWLLIPAPAPG.....RRCTKGGVSYPRAEVAETGA 459

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 97014 seqs, 13122538 residues

Total number of hits satisfying chosen parameters: 97014

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA New:*

- 1: /cgm2_6/ptodata/1/pubpaa/US08_NEW_PUB pep.*
- 2: /cgm2_6/ptodata/1/pubpaa/US06_NEW_PUB pep.*
- 3: /cgm2_6/ptodata/1/pubpaa/US07_NEW_PUB pep.*
- 4: /cgm2_6/ptodata/1/pubpaa/PCT_NEW_PUB pep.*
- 5: /cgm2_6/ptodata/1/pubpaa/US09_NEW_PUB pep.*
- 6: /cgm2_6/ptodata/1/pubpaa/US10_NEW_PUB pep.*
- 7: /cgm2_6/ptodata/1/pubpaa/US11_NEW_PUB pep.*
- 8: /cgm2_6/ptodata/1/pubpaa/US60_NEW_PUB pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2424	100.0	459	7	US-11-186-284-12
2	562	23.2	337	6	US-10-063-703-74
3	562	23.2	337	7	US-11-102-240-74
4	439	18.1	255	7	US-11-054-281-304
5	425.5	17.6	283	6	US-10-055-877-167
6	402.5	16.6	260	6	US-10-055-877-163
7	377	15.6	262	6	US-10-055-877-16
8	377	15.6	262	6	US-10-055-877-162
9	376	15.5	262	6	US-10-055-877-18
10	362.5	15.0	261	6	US-10-055-877-165
11	359.5	14.8	261	6	US-10-055-877-164
12	355	14.6	421	6	US-10-673-781-1
13	354	14.6	260	6	US-10-055-877-166
14	325	13.4	1445	7	US-11-169-041-181
15	304.5	12.6	374	7	US-11-097-728-4
16	304.5	12.6	2314	7	US-11-097-728-2
17	304.5	12.6	2353	7	US-11-097-728-6
18	300	12.4	312	7	US-11-054-281-20
19	300	12.4	312	7	US-11-054-281-86
20	292.5	12.1	266	7	US-11-054-281-88
21	290.5	12.0	294	7	US-11-054-281-87
22	290	12.0	262	7	US-11-054-281-89
23	277	11.4	328	6	US-10-131-826A-34
24	263.5	10.9	328	6	US-10-131-826A-326
25	245	10.1	252	6	US-10-467-657-276

Sequence 4008, Ap
Sequence 90, Appl
Sequence 17, Appl
Sequence 20, Appl
Sequence 18, Appl
Sequence 16, Appl
Sequence 6, Appl
Sequence 2, Appl
Sequence 37, Appl
Sequence 2, Appl
Sequence 2, Appl
Sequence 4, Appl
Sequence 786, App
Sequence 5, Appl
Sequence 444, App
Sequence 442, App

ALIGNMENTS

RESULT 1

US-11-186-284-12
; Sequence 12, Application US/11186284
; Publication No. US20050266493A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Berger, Allison
; APPLICANT: Guillemette, Tracy L.
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Thibodeau, Stephen N.
; APPLICANT: Burgart, Lawrence J.
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; FILE REFERENCE: MEMO1-029P2RNM
; CURRENT APPLICATION NUMBER: US/11/186,284
; CURRENT FILING DATE: 2005-07-21
; PRIOR APPLICATION NUMBER: US/10/301,822
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 60/339,971
; PRIOR FILING DATE: 2001-12-10
; PRIOR APPLICATION NUMBER: US 60/361,978
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/381,988
; PRIOR FILING DATE: 2002-05-20
; NUMBER OF SEQ ID NOS: 228
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 459
; TYPE: PRT
; ORGANISM: Homo Sapiens
; US-11-186-284-12

Query Match 100.0%; Score 2424; DB 7; Length 459;
Best Local Similarity 100.0%; Pred. No. 4.2e-178;
Matches 459; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAPLCPSPWLLIPAPAPGLTVQLLSLLILMPVHPORLPRMOEDSPILGGSSGSDPL 60
Db 1 MAPLCPSPWLLIPAPAPGLTVQLLSLLILMPVHPORLPRMOEDSPILGGSSGSDPL 60
QY 61 GEEDLPSEDSPREDDPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLP 120
Db 61 GEEDLPSEDSPREDDPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLP 120
QY 121 DPQEPQNNHRDKEGDDQSHWRYGGDPFVPSACAGRFQSPVDIRPQAAFCFALRPL 180

Db 121 DPEPQNNARHDEGGDDQSHRWYGGDPPWPRVSPACAGRFQSPVDIRPQLAAFCPALRPL 180
QY 181 ELGLGFOLPPLPELRNNNGHSVQLTLPGLLEMALGPGREYRALQLHLHWGAAGRPGSEHT 240
Db 181 ELGLGFOLPPLPELRNNNGHSVQLTLPGLLEMALGPGREYRALQLHLHWGAAGRPGSEHT 240
QY 241 VEGHFRPAETHVHLSTAFARVDEALGRPGGLAVLAFLAFLAEEGPEENSAYEQLLSRLEEIA 300
Db 241 VEGHFRPAETHVHLSTAFARVDEALGRPGGLAVLAFLAFLAEEGPEENSAYEQLLSRLEEIA 300
QY 301 EGESETQVGLDTSALLPSDFSRYFQYEGSLTTPPCAQGVITVFNQTVMLSAKQLHTLS 360
Db 301 EGESETQVGLDTSALLPSDFSRYFQYEGSLTTPPCAQGVITVFNQTVMLSAKQLHTLS 360
QY 361 DTGLWFGDSSLQLNFRATQPLNGRVIEASPPAGVDSPPRAAEPVQLNSCLAAGDILALVF 420
Db 361 DTGLWFGDSSLQLNFRATQPLNGRVIEASPPAGVDSPPRAAEPVQLNSCLAAGDILALVF 420
QY 421 GLLFAVTSVAFVLQVRRQHRRTGKGVSYRPAEVAETGA 459
Db 421 GLLFAVTSVAFVLQVRRQHRRTGKGVSYRPAEVAETGA 459
RESULT 2
US-10-063-703-74
; Sequence 74, Application US/10063703
; Publication No. US20060008901A1
; GENERAL INFORMATION:
; APPLICANT: Eaton, Dan L.
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3230R1C1
; CURRENT APPLICATION NUMBER: US/10/063, 703
; CURRENT FILING DATE: 2002-05-08
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 170
; SEQ ID NO 74
; LENGTH: 337
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-063-703-74
Query Match 23.2%; Score 562; DB 6; Length 337;
Best Local Similarity 38.8%; Pred. No. 5.9e-36;
Matches 128; Conservative 47; Mismatches 129; Indels 26; Gaps 7;
QY 136 DQOSHWRV---GGDPPWPRVSPACAGRFQSPVDIRPQLAAFCPALRPLELLGFPOLPLPE 192
Db 17 DGGQHWTYEGPHGDHWPASYPECGNNAQSPIDIQTDSVTFDDLPALQPHGYDQGTGP 76
QY 193 LRLRNNNGHSVQLTLPGLLEMALGPGREYRALQLHLHWGAAGRP-GSEHTVEGHRFP AEIH 251
Db 77 LDLLHNNGHTVQLSLPSTLYLG-GLPRKYVAAQLHLHWGKSPGSGSEHQINSEATFAELH 135
QY 252 VVHL-STAFARVDEALGRPGGLAVLAFLAFLAEEGPEENSAYEQLLSRLEEIAEEGSETQVPG 310
Db 136 IVHYDSYDSLSEAAERPGGLAVLGLIEVGTEKTAIYEHILSHLHEVRHKDQKTSVPP 195
QY 311 LDISALLPSDFSRYFQYEGSLTTPPCAQGVITVFNQTVMLSAKQLHTLSDTLWGP GDSR 370
Db 196 FNLRELLPKLQGYFRNGSLTTPPCYQSVLWTVFYRRSQISMEQLEKLGQTLFSTEEEP 255
QY 371 LQL---NFRATQPLNGRVIEASPPAGVDSPPRAAEPVQLNSCLAAGDILALVFGLLFA-- 425
Db 136 IVHYDSYDSLSEAAERPGGLAVLGLIEVGTEKTAIYEHILSHLHEVRHKDQKTSVPP 195
QY 311 LDISALLPSDFSRYFQYEGSLTTPPCAQGVITVFNQTVMLSAKQLHTLSDTLWGP GDSR 370
Db 196 FNLRELLPKLQGYFRNGSLTTPPCYQSVLWTVFYRRSQISMEQLEKLGQTLFSTEEEP 255
QY 371 LQL---NFRATQPLNGRVIEASPPAGVDSPPRAAEPVQLNSCLAAGDILALVFGLLFA-- 425

Db 256 SKLLVQNYRALQPLNORMVFPASF-----IQAGSSYTTGEMLSGLGVILVGCL 302
QY 426 --VTSVAFVLQVRRQHRRTGKGVSYRPAE 453
Db 303 CULLAVYFIARKIRKRLNKRKSVVFTSAQ 332
RESULT 3
US-11-102-240-74
; Sequence 74, Application US/11102240
; Publication No. US20050260647A1
; GENERAL INFORMATION:
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: ANTIBODIES TO POLYPEPTIDES ENCODED BY A NUCLEIC ACID UNDEREXPRESSI
; TITLE OF INVENTION: ESOPHAGEAL TUMOR
; FILE REFERENCE: P3230R1C106C
; CURRENT APPLICATION NUMBER: US/11/102,240
; CURRENT FILING DATE: 2005-04-08
; PRIOR APPLICATION NUMBER: 10/063662
; PRIOR FILING DATE: 2002-05-07
; PRIOR APPLICATION NUMBER: 10/006867
; PRIOR FILING DATE: 2001-12-06
; PRIOR APPLICATION NUMBER: PCT/US00/23328
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: 60/170262
; PRIOR FILING DATE: 199-12-09
; NUMBER OF SEQ ID NOS: 170
; SEQ ID NO 74
; LENGTH: 337
; TYPE: PRT
; ORGANISM: Homo Sapien
US-11-102-240-74
Query Match 23.2%; Score 562; DB 7; Length 337;
Best Local Similarity 38.8%; Pred. No. 5.9e-36;
Matches 128; Conservative 47; Mismatches 129; Indels 26; Gaps 7;
QY 136 DQOSHWRV---GGDPPWPRVSPACAGRFQSPVDIRPQLAAFCPALRPLELLGFPOLPLPE 192
Db 17 DGGQHWTYEGPHGDHWPASYPECGNNAQSPIDIQTDSVTFDDLPALQPHGYDQGTGP 76
QY 193 LRLRNNNGHSVQLTLPGLLEMALGPGREYRALQLHLHWGAAGRP-GSEHTVEGHRFP AEIH 251
Db 77 LDLLHNNGHTVQLSLPSTLYLG-GLPRKYVAAQLHLHWGKSPGSGSEHQINSEATFAELH 135
QY 252 VVHL-STAFARVDEALGRPGGLAVLAFLAFLAEEGPEENSAYEQLLSRLEEIAEEGSETQVPG 310
Db 136 IVHYDSYDSLSEAAERPGGLAVLGLIEVGTEKTAIYEHILSHLHEVRHKDQKTSVPP 195
QY 311 LDISALLPSDFSRYFQYEGSLTTPPCAQGVITVFNQTVMLSAKQLHTLSDTLWGP GDSR 370
Db 196 FNLRELLPKLQGYFRNGSLTTPPCYQSVLWTVFYRRSQISMEQLEKLGQTLFSTEEEP 255
QY 371 LQL---NFRATQPLNGRVIEASPPAGVDSPPRAAEPVQLNSCLAAGDILALVFGLLFA-- 425
Db 256 SKLLVQNYRALQPLNORMVFPASF-----IQAGSSYTTGEMLSGLGVILVGCL 302
QY 426 --VTSVAFVLQVRRQHRRTGKGVSYRPAE 453
Db 303 CULLAVYFIARKIRKRLNKRKSVVFTSAQ 332
RESULT 4
US-11-054-281-304
; Sequence 304, Application US/11054281
; Publication No. US20060013813A1
; GENERAL INFORMATION:
; APPLICANT: Mezes et al.
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same

```

; APPLICANT: Andrew, David
; APPLICANT: Mezas, Peter
; APPLICANT: Patturajan, Meera
; APPLICANT: Burgess, Catherine
; APPLICANT: Eisen, Andrew
; APPLICANT: Wolenc, Adam
; APPLICANT: Baumgartner, Jason
; APPLICANT: Shinkets, Richard
; APPLICANT: Gusev, Vladimir
; APPLICANT: Vernet, Corine
; APPLICANT: Taupier Jr., Raymond
; APPLICANT: Pena, Carol
; APPLICANT: Shenoy, Suresh
; APPLICANT: Li, Li
; APPLICANT: Casman, Stacie
; APPLICANT: Boldog, Ferenc
; TITLE OF INVENTION: Novel Polypeptides and Nucleic Acids Encoded Thereby
; FILE REFERENCE: 21402-251
; CURRENT APPLICATION NUMBER: US/10/055,877
; CURRENT FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: 60/262,892
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: 60/263,598
; PRIOR FILING DATE: 2001-01-23
; PRIOR APPLICATION NUMBER: 60/263,799
; PRIOR FILING DATE: 2001-01-24
; PRIOR APPLICATION NUMBER: 60/264,117
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 60/264,139
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 60/264,478
; PRIOR FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 60/263,351
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/272,870
; PRIOR FILING DATE: 2001-03-02
; PRIOR APPLICATION NUMBER: 60/275,990
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/275,927
; PRIOR FILING DATE: 2001-03-14
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 512
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 167
; LENGTH: 283
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: carb_anhydase
; OTHER INFORMATION: domain consensus sequence
; US-10-055-877-167

Query Match 17.6%; Score 425.5; DB 6; Length 283;
Best Local Similarity 36.1%; Pred. No. 1.3e-25;
Matches 105; Conservative 42; Mismatches 103; Indels 41; Gaps 11

Qy 124 EPQNNHRDKGDDQSHWRYGGDPPRPVSPA-----CAGRFQSPVDIRPOLAAFCPALR 178
      ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
Db 10 EHSNNAH-----VLMHKLYPIANGCNCQGERQSPFINIQTKEAKYDPSLK 53

Qy 179 PLELLGFLQPLPELRLRNNGHSVOLTLPPGLEMAL---GP---GREYRALQLHLHWGAA 232
      ||| : : : : : || : : : : : || : : : : : || : : : : : || : : : : :
Db 54 PLS-LSYDAATAKEFEIVNNGHSFQVEFDDSDDKSVLSGGDLPHAGHPYRLKQFHFWGGA 112
      || : : : : : || : : : : : || : : : : : || : : : : : || : : : : :

Qy 233 GR--PGSEHTVEGRFPFAEIHVHL--STAFARVDEALGRPGGLAVLAALFEEGP--EENSA 288
      || : : : : : || : : : : : || : : : : : || : : : : : || : : : : :
Db 113 SSDQGSSEHVDGKKYAAELHLVHWNSTKGYSEKAVSKPDGLAVGLGVFLKVDGYQENPG 172
      || : : : : : || : : : : : || : : : : : || : : : : : || : : : : :

Qy 289 YEQLLSLEIAEBSGSTQVGLDIALSPDFSR-YFOYEGSLTTPPCAQGVITWVFNQ 347
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 173 LQKVVDALSSIKTKGSGATFNFPSTLLPSEKLRDYNWTYPGSLTTPPLTESVTWVLKE 232
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Qy 348 TVMSAKQLHTLSDTLWG-----PG-DSFLQLNFRATQPLNGRVEIASF 390

```


1


```
Query Match 15.0%; Score 362.5; DB 6; Length 261;
Best Local Similarity 32.6%; Pred. No. 7.5e-21;
Matches 87; Conservative 46; Mismatches 111; Indels 23; Gaps 8;

QY 132 DKEGDDQSHWRYGDPWPVRVSPACAGRFQSPVDIRPQLAFCPALRPLELLGFLQPLP 191
DB 10 DKNQPEQ-----WSKLYPIANGNNSPVDIKTSETKHDTSLKPISV---SYNPAT 56

QY 192 ELRLNNGHVSQVLTLPGLLEMAL---GP-GREYRALQLHLHWGAAGRPGESEHTVEGHRFP 247
DB 57 AKIINVGHSFHVTFEDNDRSVLKGGLPSDYLRFQHFHMGSTNEHSGSEHTVDGVKYS 116

QY 248 AEIHVHLSTA-FARVDEALGRGGLAVLAAPLEEGPEENSAYEQLLSRLEETAEBSSET 306
DB 117 AEHLTHWNSAKYSLSLAASAKADGLAVIGVLMKVG-EANPKLOKILDALQAIKTKGRA 175

QY 307 QVPGDLISALLPSDFSRYFOYEGSLTTPPCAQQGVITVFNQTVMLSAKQL---HTLSDTL 363
DB 176 PFTNFDSTLLPSLD-FWTYPGSLTHPPLYESVTWICKESISVSSEQLAQFRSLLSNV 234

QY 364 WPGDSDRLQNLFRATQPLNGRVIEASF 390
DB 235 EGDNAVPMQHNRRPTQLKGRTRASF 261

RESULT 11
US-10-055-877-164
; Sequence 164, Application US/10055877
; Publication No. US20050288241A1
; GENERAL INFORMATION:
; APPLICANT: Decristofaro, Marc
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Miller, Charles
; APPLICANT: Tchernev, Velizar
; APPLICANT: Zhong, Mei
; APPLICANT: Anderson, David
; APPLICANT: Ballinger, Robert
; APPLICANT: Gerlach, Valerie
; APPLICANT: Spytek, Kimberly
; APPLICANT: Ratelli, Luca
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Guo, Xiaojia
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Andrew, David
; APPLICANT: Mezes, Peter
; APPLICANT: Patturajan, Meera
; APPLICANT: Burgess, Catherine
; APPLICANT: Eisen, Andrew
; APPLICANT: Wolenc, Adam
; APPLICANT: Baumgartner, Jason
; APPLICANT: Shimkets, Richard
; APPLICANT: Gusev, Vladimir
; APPLICANT: Vernet, Corine
; APPLICANT: Taupier Jr., Raymond
; APPLICANT: Pena, Carol
; APPLICANT: Shenoy, Suresh
; APPLICANT: Li, Li
; APPLICANT: Casman, Stacie
; APPLICANT: Boldog, Ference
; TITLE OF INVENTION: Novel Polypeptides and Nucleic Acids Encoded Thereby
; FILE REFERENCE: 21402-251
; CURRENT APPLICATION NUMBER: US/10/055,877
; CURRENT FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: 60/262,892
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: 60/263,598
; PRIOR FILING DATE: 2001-01-23
; PRIOR APPLICATION NUMBER: 60/263,799
; PRIOR FILING DATE: 2001-01-24
; PRIOR APPLICATION NUMBER: 60/264,117
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 60/264,139

Query Match 14.8%; Score 359.5; DB 6; Length 261;
Best Local Similarity 32.6%; Pred. No. 1.3e-20;
Matches 87; Conservative 45; Mismatches 112; Indels 23; Gaps 8;

QY 132 DKEGDDQSHWRYGDPWPVRVSPACAGRFQSPVDIRPQLAFCPALRPLELLGFLQPLP 191
DB 10 DKNQPEQ-----WSKLYPIANGNNSPVDIKTSETKHDTSLKPISV---SYNPAT 56

QY 192 ELRLNNGHVSQVLTLPGLLEMAL---GP-GREYRALQLHLHWGAAGRPGESEHTVEGHRFP 247
DB 57 AKIINVGHSFHVTFEDNDRSVLKGGLPSDYLRFQHFHMGSTNEHSGSEHTVDGVKYS 116

QY 248 AEIHVHLSTA-FARVDEALGRGGLAVLAAPLEEGPEENSAYEQLLSRLEETAEBSSET 306
DB 117 AEHLTHWNSAKYSLSLAASAKADGLAVIGVLMKVG-EANPKLOKILDALQAIKTKGRA 175

QY 307 QVPGDLISALLPSDFSRYFOYEGSLTTPPCAQQGVITVFNQTVMLSAKQL---HTLSDTL 363
DB 176 PFTNFDSTLLPSLD-FWTYPGSLTHPPLYESVTWICKESISVSSEQLAQFRSLLSNV 234

QY 364 WPGDSDRLQNLFRATQPLNGRVIEASF 390
DB 235 EGDNAVPMQHNRRPTQLKGRTRASF 261

RESULT 12
US-10-673-781-1
; Sequence 1, Application US/10673781
; Publication No. US20050260689A1
; GENERAL INFORMATION:
; APPLICANT: Qirwei Shi
; TITLE OF INVENTION: Differential Immunoassay
; FILE REFERENCE: 1112-1-080N
; CURRENT APPLICATION NUMBER: US/10/673,781
; CURRENT FILING DATE: 2003-09-29
; PRIOR APPLICATION NUMBER: US/09/938,270B
; PRIOR FILING DATE: 2001-08-23
; PRIOR APPLICATION NUMBER: US 60/227,536
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/292,497
; PRIOR FILING DATE: 2001-05-21
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 421
; TYPE: PRT
; ORGANISM: homosapien
US-10-673-781-1

Query Match 14.6%; Score 355; DB 6; Length 421;
Best Local Similarity 30.1%; Pred. No. 5.2e-20;
Matches 94; Conservative 55; Mismatches 137; Indels 26; Gaps 9;
```


Db	53	KASGDP-----YWAYS	GAYGPEHWTSSVSCGRHOSP	IDILQYARV	95
Qy	174	CPALRPDELLG	FQLPPLPELRLNRNGHSVQLTLP-----	PGLEMALGPGREYRALQLH	226
Db	96	GEEVQELQLQDGF	DNESNKTMTKGTVAILLKDDYFVSGAGL-----	PCR-PKAEKV	149
Qy	227	LHWG-AACRPGSEHTV	GCHRPALIHVVHLS-TAFARVDEALG	PPGGLAVLAALFLEEGPE	284
Db	150	PHWGHSNGSAGSEHS	INGRRPFVEVQFIFFYPDDPDSFQT	AISENRIIGMAIFQVQSPR	209
Qy	285	ENSAEQQLARLEB	IAEGSTVQVGLDISALLPSDFSRYFYQ	EGSLTPPCAQGVITWV	344
Db	210	DNSALDPIIHGLKGV	WHKEKTFLLDPFVRLDLLPASLG	SVYRYTGSLLTPPCSEIVEWV	269
Qy	345	FNQTYMLSAKOLH	LTSDTLWGPCDGR-----LQ	LNFRATQPLNGRVIEASP	390
Db	270	FRRPVPISYHOLEA	FYSIFTTQEQDHVKSVEYLRNFRPQ	QLRHDVVVSKSAVRDSWNHD	329
Qy	391	-----PAGVDSS	PRAAEP	403	
Db	330	MTDFLENPLGT	EASKVCSSP	349	

RESULT 15

US-11-097-728-4
; Sequence 4, Application US/11097728
: Publication No. US20050260132A1

```

1  / FUNDATION NO: 00200302001J2K4
2  / GENERAL INFORMATION:
3  / APPLICANT: Erik Foehr
4  / APPLICANT: Sabine Muller
5  / APPLICANT: Daniel J. Chin
6  / APPLICANT: Mirella Gonzalez-Zulueta
7  / TITLE OF INVENTION: MONOCLONAL ANTIBODIES DIRECTED TO RECEPTOR PROTEIN
8  / TITLE OF INVENTION: TYROSINE PHOSPHATASE ZETA
9  / FILE REFERENCE: AGYT-006CIP
10 / CURRENT APPLICATION NUMBER: US/11/097,728
11 / CURRENT FILING DATE: 2005-03-31
12 / PRIOR APPLICATION NUMBER: 10/652,981
13 / PRIOR FILING DATE: 2003-08-28
14 / NUMBER OF SEQ ID NOS: 14
15 / SOFTWARE: FaatSEQ for Windows Version 4.0
16 / SEQ ID NO 4
17 / LENGTH: 374
18 / TYPE: PRT
19 / ORGANISM: Homo sapiens
20 / US-11-097-728-4

```

		Query Match	12.6%; Score 304.5; DB 7; Length 374;
		Best Local Similarity	28.0%; Pred. No. 3.2e-16;
		Matches	87; Conservative 50; Mismatches 149; Indels 25; Gaps 8;
Qy	141	WRYGG---DPPPRVSPACAGRQSPVDIRPOLAACPALRPLELLGFOQLPPLPELRN	197
Db	38	WSYTGA LKNQWGKKYPTCNPSKQSPINIDEDLTQVNVLKCKLFGQWDKTSLENTFIHN	97
Qy	198	NHGSVQTLTPGLMALGPGR-EYRALQLHLHWCAA--GRPGESEHTVEGHFPAEIIHVH	254
Db	98	TGKTVEINLTNDYRVSGVSEMFKASKITFWHGKNCNMSSDGSHSLEGOKFPDEMQLYC	157
Qy	255	L-STAFARVDEALGRPGCLAVLAFLBEGEPENSAVEBQLLSRLBEIAEBESQETOPVGDI	313
Db	158	FDADRFSSFBEAVNGKGKJLRALSILFEVGTENLDLFRAIDGVESVSRRPGKAALDPFIL	217
Qy	314	SALLPSDFSRYFYQYEGSLTTPPCAQQGIWTVFNOTVMLSAKQLHTLSDTLMGPGD----	368
Db	218	LNLPLNSTDKYYIYNGSLTSPCTCDTDVIVFKDTVISISSQLAVFCEVLTMQOSGYML	277
Qy	369	-SRQLMFRATQPLNGRVIASPF-----PAGVDSPPR--AAPPVQNINSCLAGDILAL	418
Db	278	MDYLQNNFRECOYKFYSRQVFSSTYGKEIEHAVCSSEPENVQADPENVTSLTVTWERPRV	337
Qy	419	VFGLL---FAV 426	

THIS PAGE BLANK (USPTO)